TESI DI DOTTORATO

UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II

DIPARTIMENTO DI INGEGNERIA ELETTRONICA
E DELLE TELECOMUNICAZIONI

DOTTORATO DI RICERCA IN
INGEGNERIA ELETTRONICA E DELLE
TELECOMUNICAZIONI

A Fast and Non Redundant Approach to the Analysis of Large Scatterers

MARCO SANTOJANNI

Il Coordinatore del Corso
Ch.mo Prof. Giovanni Poggi

Il Tutore
Ch.mo Prof. Giuseppe D'Elia

Anno accademico 2005–2006

Index

Chapter 1

Introduction

The analysis of the electromagnetic scattering plays a key role in systems, applications, and services in present and future Information and Communication Technology (ICT), as well as in fields like medical diagnosis and telemedicine, earth observation and monitoring, non invasive electromagnetic sensing, and so on. In these areas, electromagnetic field engineering involves the design and performance prediction of devices, like antennas and radiofrequency (RF) circuits, that are designed to produce and radiate fields, the prediction of wave scattering (as in radar, non invasive and remote-sensing systems), as well as the prediction of adverse effects of unwanted fields on systems and subsystems, as in typical Electromagnetic Compatibility (EMC) endeavours. In these applications, the use of numerical methods, able to reduce the prototyping and/or testing time, has grown in recent years following the increasing availability of computing power, leading the computational electromagnetics to significant achievements. However, the interest in the analysis of large and complex problems, jointly with the trend toward the use of higher and higher frequencies, has to be taken into account. Despite the high performances that computers can currently offer, the computational effort needed in a variety of situations of definite industrial or scientific interest still remains a really expensive, or even prohibitive, task.

Both large electrical size and geometrical complexity contributes to the scale of the numerical problem. Classical examples of large scale problem is the evaluation of the radar cross section of large objects, the estimation of radio coverage in cellular or other RF communication systems, the fast prediction of the scattering of bodies in inverse problems (e.g. in the Ground Penetrating Radar) for the detection of buried objects, the full analysis of large antennas.

Many approaches have been proposed by the researchers to solve this problem.

Asymptotical methods allow to analyse very large structures with low numerical load, but their effectiveness is strongly dependant on the problem in use: complex geometries cannot be studied in a satisfactory way, multiple interaction cannot be taken into account.
The answer to this kind of problem should be a "full wave" method, flexible and able to numerically solve a wide class of problems.
The numerical full wave method approaches belong to two main categories: those based on a Differential Formulation (DF) and those based on an Integral Formulation (IF). In both cases, the aim is to express the problem in form of a linear system, defined (and to be solved) in a finite dimensional space. Among the DF-based methods, are the Finite Difference Time Domain method (FDTD) and the Finite Element Method (FEM), this latter in the frequency domain. The most adopted IF-based technique goes by the name of Method of Moments (MoM).
DF and IF techniques have different characteristics: a comparison is possible between FEM and MoM, both operating in frequency domain. In FEM techniques, the differential equations that support the EM field are directly discretized in the considered domain: as a consequence, a lattice has to be defined in the whole volume, and the unknowns are distributed in 3 dimensions. The resulting linear system is made by many equations, each involving a few number of unknowns, describing a large but strongly sparse matrix. As a consequence, a further complicacy is that an infinite domain couldn't be directly considered. For example, in the study of a scatterer in free space has to be approximated with a truncated domain with absorbing condition on its boundaries.
In MoM techniques, the unknowns are distributed on the boundary surfaces between different media, so that their number grows with the square of the linear dimensions instead on the cube, and play the role of equivalent currents. The resulting linear system has a physical meaning, because relates such currents to the radiated fields, and is defined by a so-called Impedance Matrix, which is quite smaller than the FEM one, even if it is dense rather than sparse. At difference with the FEM, there is no need to impose additional absorbing conditions to take into account the radiation condition at infinite, allowing to easily analyze scattering systems in free space. Moreover, considerations on the physical interpretation of the linear systems

allow to adopt some approximation that can result in reducing the computational complexity of the solution. As a matter of fact, the MoM techniques are the most employed ones for large scatterers in linear media.
Let us now investigate the issue of assessing the size of the problem.

1.1 Problem statements

The classical scattering problem can be stated as follows: given a scattering system (for example, the antenna structure) and an excitation (a set of primary sources or an incident field), evaluate the total field (sum of incident and scattered fields) on a given observation domain.
The IF-based methods, like the Moment Method, are all based on the following steps:

- the relationship between a generic distribution of equivalent sources, placed on the boundary surface(s), and the radiated field is expressed;
- an expansion of these equivalent sources is found by enforcing the boundary condition;
- the field radiated by the equivalent current distribution is evaluated on the observation domain, and gives the estimate of the scattered field.

The first two points consist in expressing and solving the source equation:

$$GJ=F_0 \qquad (1.1)$$

where J is the induced current density, defined on the boundary domain, G is the radiation operator and depends on the adopted formulation (more details are given in the following chapters), F_0 the field due to the primary excitation on the boundary domain..
The third point amounts to evaluating the radiation integral:

$$E=G_0J \qquad (1.2)$$

where J is the previously found current distribution, E is the scattered field in the given observation domain, G_0 is the appropriate radiation operator. Obviously, to be practically implemented, the current expansion must be approximated by a finite number of terms.

In MoM approach, the problem is solved by projecting the equation (1.1) on a suitable finite dimensional space, thus finding a solution J in terms of an expansion in basis functions. Obviously, the space dimension grows with the electrical size of the boundary surfaces involved in the problem. Being N the finite dimension of such space, the problem is turned to the solution of a system of linear equations and is expressed in the form of a matrix-by-vector equation:

$$\underline{\underline{G}}\,\underline{J} = \underline{F}_0 \qquad (1.3)$$

where the N-by-N matrix $\underline{\underline{G}}$ is called impedance matrix, and the vectors $\underline{J}$ and $\underline{F_0}$ express the expansion coefficients of the induced current and the forcing field, respectively.

The application of the M.o.M.-like techniques is characterized by the following main costs:

- a memory occupation for the storage of the elements of the matrix describing the radiating operator G (impedance matrix). This occupancy is of the order N^2;
- a computational complexity of order $O(N^2)$ for the evaluation of the N^2 elements of the impedance matrix;
- a computational complexity of order $O(N^3)$ for the inversion of the impedance matrix (i.e. for the solution of the linear system)

Accordingly, increasing the electrical size of the scattering system, increase the computational costs more than linearly. In particular, the effort required for the inversion of the impedance matrix rapidly becomes unsustainable.

The main approaches introduced to keep down the computational cost of this kind of techniques essentially attempt to reduce the number of unknowns as well as to introduce appropriate algorithms reducing the memory requirements and the computational complexity of the problem.

While in the first attempt the choice of a suitable sets of expansion functions is involved, the latter strategy relies on the consideration that the MoM impedance matrix, even if dense, contains highly

redundant information, that can be conveniently handled by means of fast schemes able to reduce the storage and to speed up the inversion process.
Before discussing the various approaches proposed for the evaluation of the electromagnetic scattering from electrically large structures, in the following the estimation of the minimum size and computational cost of the of the scattering problem is addressed

1.2 Assessing the size of the problem

The key question to answer before trying to reduce the computational effort related to these issues is whether it exists a lower bound to the numerical complexity for a given problem, analogous to the famous result of Shannon for information and communication. If this value can be determined, it will be possible to ask whether it exists an algorithm able to solve the EM scattering problem with the corresponding minimal effort.
The first problem is to determine the best approximation of the radiation operator among the set of finite rank operators.
The compactness of the operator G_0 allows to approximate it, with an arbitrary error, by means of a finite rank operator [1]. The best approximation of rank n for G_0, that we can indicate with G_0^n, is obtained by truncation the SVD decomposition to the first n terms. For the radiation operator G_0, the SVD is:

$$G_0 = \sum_{i=0}^{\infty} \left| e_i^0 > \sigma_i < j_i^0 \right| \qquad (1.4)$$

where σ_i are the singular values, $\left|e_i^0\right.>$ and $<\left. j_i^0\right|$ are the corresponding left and right singular vectors, providing an orthonormal basis in the output and input spaces, respectively. The singular values σ_i form a non-increasing sequence going to zero as $n \to \infty$, and we have:

$$\frac{\sigma_n}{\sigma_0} = \frac{\left\| G_0^n - G_0 \right\|}{\left\| G_0 \right\|} \tag{1.5}$$

where $\|\cdot\|$ is the operator norm. This equation allows to evaluate the number of significant terms of the SVD, provided that the singular values can be calculated analytically. Since this is possible only for particular, simple scattering geometries, a general approach for the evaluation of the rank of the finite dimensional operator approximating the radiation operator is proposed by exploiting the properties of the kernel of G_0. In [2] it is shown that for large scattering systems, the sequence of singular values has a step-like behaviour (see fig. 1.1): it decreases slowly until a critical value of n, say n', is reached, then starts to decay very rapidly. This implies that, when substituting the radiating operator with a finite dimensional one, in order to obtain reasonable values for the approximation error, the needed rank N_0 can be chosen slightly larger than n', and it will be substantially independent on the required precision.
The so found value N_0 can be considered the effective dimension of the scattering problem, and can be estimated as [3]:

$$N_0 \cong 2\frac{Area(\Sigma)}{(\lambda/2)^2} \tag{1.6}$$

Σ being the surface of the smallest convex envelope including the scattering system.

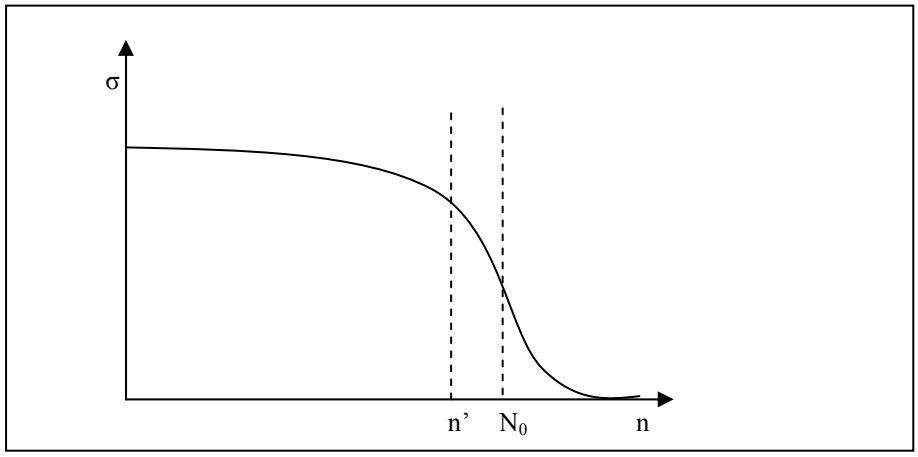


Figure 1.1 – Behaviour of the sequence of the singular values

The study of the other operator implied in the problem, G, is not as easy as for the operator G_0. In fact, the inverse operator G^{-1} is unbounded (its role is to relate the equivalent sources on the scatterer surface with the field in the same domain), so that it doesn't exist any finite rank approximation that allows the evaluation of a generic current distribution J with a bounded error. This fact seems to neutralize the attempt to define the size of the problem. However, we have to remember that we're not in a completely general situation: in fact, we suppose that the incident field is due to uniformly bounded sources external to the antenna/scatterer support. The compactness of the radiation operator implies that the set of any possible radiated field is compact, so that the effective field can be approximated with a bounded error by the elements of a finite dimensional subspace. According to [2], if the primary sources are far enough from the scattering system (at least some wavelengths), the dimension of the subspace is essentially equal to N_0. Thus, the operator G can be approximated by $G_{N0}=P_{N0}G$, P_{N0} being the orthogonal projector onto the subspace of the incident fields [1]. In this way, we can say that the dimension of the whole problem is N_0. This is in accordance with the results shown in [4], where it is found that the set of the first N_0

characteristic fields provides an accurate representation of the scattered field.

Provided that it is possible to express the operator G by means of a low rank approximation, say G_{N0}, we want to invert the operator, and solve the source equation, with the minimum numerical effort. The ideal lower limit for the complexity is $O(N_0)$, which we should have if the SVD form of the problem were available in advance. This is not our case, since what we have in practice is just the impedance matrix provided by the IF-based methods. The size of this square matrix, say N, depends on the implementation of the method, and in general it does not correspond with N_0. Moreover, the matrix is fully populated, and the singular value decomposition would require itself a computational complexity of the order $O(N^3)$. Analogously, the direct inversion of the matrix has to be carried out by means of algorithms as the LU decomposition or the Gaussian elimination, also requiring a cubic computational complexity.

1.3 Reduction of the computational costs

The reduction of the computational costs involves the two steps:

- finding a practical method to obtain a minimal rank numerical expression of the radiation operator
- finding a scheme to save memory storage and speed up the inversion process.

The first point concerns the possibilities of studying the problem by means of a number of unknowns equal to N_0. With reference to the MoM-based techniques, the problem consists in using a minimal set of basis functions able to express the scattered fields with a bounded error. From this point of view, for large metallic scatterers, particularly attractive appears the Method of Auxiliary Source, which will be described and implemented in the following.

Regarding the second point, many ways have been suggested and covered. The first consideration is that the solution of a linear system can be found by means of an iterative process, with a bounded approximation, instead that in direct way. The iterative algorithms (CG, BIGC, GMRES…) [5,6] essentially involve a matrix by vector product at each step. Assuming that a bounded error solution is

reached within a number of steps independent on the size of the problem, the computation effort of the inversion is equal to that of the matrix-by-vector product, i.e. $O(N_0^2)$. As seen, even if reduced, it is still far from the ideal case $O(N_0)$. The problem now is to minimize the effort needed to perform the multiplication of the impedance matrix by the tentative solution vector at each step. A class of techniques starts from the consideration that the impedance matrix contains high redundantly information, so that it can be expressed (and consequently, stored and inverted) in by exploiting a reduced redundancy scheme, with significant cost savings. In this way, a reduction of the computational complexity up to $O(N\log N)$ is achievable [7]. Among these techniques, we will focus our attention on the so called Fast Multipole Method.

Chapter 2

Overview of Fast Methods

The research on effective methods for the full-wave analysis of the electromagnetic scattering has explored two major pathways: the first aims to represent and study the scattering problem with a minimal set of variables; the second, tries to speed up the solution process of the linear system arising from the integral formulation.
The issue concerning the reduction of the unknowns in MoM-like techniques has been often addressed relying on the choice of minimal sets of basis functions, tailored on the problem under study and able to express the equivalent currents by means of very few terms.
In 1994, Boag and Mittra proposed the Complex Multipole Beam Approach (CMBA) [8]: the reduction of the MoM matrix size, for smooth surfaces, is achieved by exploiting basis functions that resemble the Gabor expansion functions.
In the CBF method [9], Characteristic Basis Functions are estimated by solving the inversion problem in restricted areas, then applied to the study of the whole scattering system.
In the Synthetic Function Expansion (SFX) approach [10-13], the scattering structure is initially subdivided in blocks, then the analysis of the single block is carried out with reduced effort. The solution are used as (synthetic) basis functions for the solution of the entire problem.
For the analysis of large metallic scatterers, a very attractive technique is the Method of Auxiliary Sources [see also chapter 3] [14-19]. In the MAS, a set of elementary point-wise sources is placed beneath the scatterer surface: it is shown that, if properly placed, very few auxiliary sources can represent the scattered field with a bounded error [20].
A branch of techniques, rather than reducing the number of unknowns, aims to speed up the solution of the linear system arising from the MoM approach, i.e. the inversion of the impedance matrix. A common

characteristic of this class of methods is the adoption of an iterative inversion algorithm (CG, BiCG, GMRES, etc…): in this way, the challenge is to expedite the evaluation of each iteration, essentially involving an (impedance)matrix-by-(tentative solution)vector product.
Since sparse matrices need fewer storage memory and can be multiplied faster than dense ones, many methods have been developed to "sparsify" the impedance matrix, by applying suitable transformations to the impedance matrix, amplifying the matrix element dynamics and cutting the smaller ones. Among these methods, particular mention is due to the Impedance Matrix Localization (IML) by Canning [21-23] and to various wavelet-based approaches, that can be classified in two branches: the Multi Resolution Methods of Moments (MRMoM) [24, 25, 26] and the Discrete Wavelet Transformation (DWT) methods [27, 28, 29].
A further way to achieve a significant reduction of memory requirements and computational complexity is the use of the Conjugate Gradient-Fast Fourier Transform technique (CGFFT), exploiting the translational invariance of the Green's operator to apply an FFT-like scheme to the matrix-by-vector product involved in the iterative solution process [30, 31, 32].
A very promising approach to the reduction of the computational costs related to the MoM goes by the name of Fast Multipole Method [see also chapter 4] [33]. The key idea under the FMM, and its further progresses, is the consideration that the off-diagonal blocks on the impedance matrix, that represent the interactions between parts of the structure far away from each other, are almost rank-deficient and can be represented in an aggregate way, reducing the information redundancy (and, as a consequence, the memory requirements) and providing a faster matrix-by-vector product scheme.
In the following, a brief review of the mentioned techniques is provided.

2.1 The Characteristic Basis Function Method

The MoM based techniques involve a domain discretization that strongly influences their performances. Usually, the MoM based techniques adopt generic kinds of basis functions, like the Rao-Wilton-Glisson (RWG) ones [34], and requires more than 10 unknowns per wavelength to give a good representation of the current density [9]. For increasing sizes of the scattering system, the size of the impedance matrix, that relates the induced current with the incident field, grows very rapidly and the problem becomes soon unmanageable.

The approach by Boag and Mittra [9] proposes the adoption of high-level expansion functions specially constructed to fit the object geometry by taking into account the physics of the problem: the Characteristic Basis Functions (CBFs). The aim is to efficiently represent the equivalent source distribution on the boundary by means of a reduced set of basis functions, each ranging over some wavelengths, thus obtaining an impedance matrix of reduced size.

The scattering geometry is subdivided in M blocks, partially overlapping at the boundaries. Then the scattering problem is solved in each isolated block, with an excitation obtained by windowing the original incident field. For the i_th block, the RWG domain decomposition is applied and the impedance matrix $Z_e^{(i)}$ accounting for the interaction within the block is explicitly evaluated. The figure 2.1 shows the positioning of the block $Z_e^{(i)}$ in an hypothetical complete impedance matrix.

In formulas, the following linear system has to be solved:

$$Z_e^{(i)} J_i^{(i)} = R^{(i)} \qquad (2.1)$$

for i=(1, 2, …, M), $R^{(i)}$ being the windowed excitation. The M solutions $J_i^{(i)}$ will constitute the primary bases for the study of the overall problem.

The next step amounts in evaluation the secondary bases, i.e. those taking into account the mutual interaction between the blocks. For each of the M blocks, M-1 secondary bases $J_k^{(i)}$ are found by solving:

$$Z_e^{(i)} J_k^{(i)} = R_k^{(i)} \tag{2.2}$$

for k=(1, 2, …, i-1, i+1, …, M), $R_k^{(i)}$ being the excitation in the i_th block due to the primary base of the k_th block. In this way, a set of M^2 basis functions (each being the aggregation of a high number of RWG bases) are found, and a reduced impedance matrix, of size M^2 by M^2, can be evaluated. In order to reduce the condition number, it is possible to orthonormalize the set of basis functions by means of the Graham-Schmidt procedure before evaluating the impedance matrix.
One of the strong points of the CBF method is the cost saving achievable when interested in multiple excitation (a very useful feature in RCS evaluations). In fact, the most computational cost of the CBF methods derives from the LU factorizations that arise from the solution of the single blocks. Since this task is independent on the excitation, such factors can be computed once, stored, and then used to solve as many excitation problems are needed with a low extra computational effort.

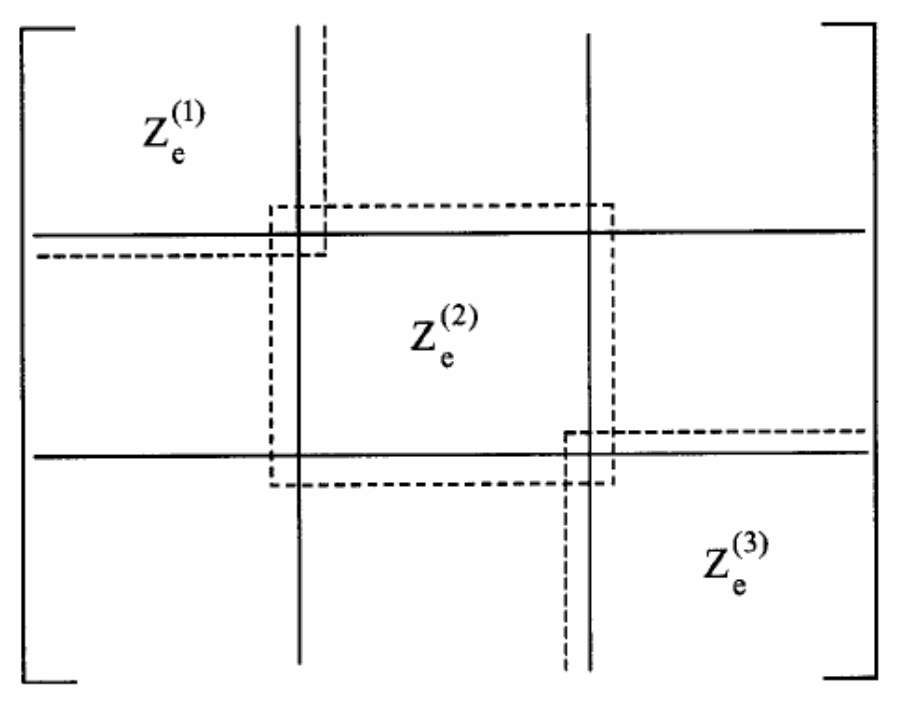


Figure 2.1– Positioning of the $Z_e^{(i)}$ blocks in the complete impedance matrix

2.2 The Complex Multipole Beam Approach

The Complex Multipole Beam approach, proposed by Boag and Mittra [8], aims to reduce the number of unknowns needed by the standard MoM techniques to study the problem of electromagnetic scattering. The key idea is to consider as basis functions the fields produced by multipoles, but shifted in the complex space. In 2D, let us consider a multipole source located in r_0 (see fig. 2.1). The z-directed electric field due to the source in the observation point $r=(\rho,\varphi)$, except for a constant factor, is given by:

$$\Psi_n(\underline{r},\underline{r}_0) = H^{(2)}_{|n|}(k\rho)e^{jn\varphi} \qquad (2.3)$$

where $H^{(2)}_{|n|}$ is the Hankel function of second kind and order n, k the propagation constant, ρ and φ are modulus and angle of the vector $\underline{r}$-$\underline{r}_0$. The time dependence $exp(j\omega t)$ is suppressed.
The source is "shifted" in the complex space by taking $\underline{r}_0 = \underline{r}_0' + j\underline{r}_0''$, $\underline{r}_0'$ and $\underline{r}_0$ being real vectors. The authors define the field of an n_th order multipole located in the complex space as:

$$\Psi_n(\underline{r},\underline{r}_0) = H^{(2)}_{|n|}(k\rho)\left(\frac{x-x_0}{\rho} + j\frac{y-y_0}{\rho}\right)^n \qquad (2.4)$$

where $\rho = \sqrt{(x-x_0)^2 + (y-y_0)^2}$.

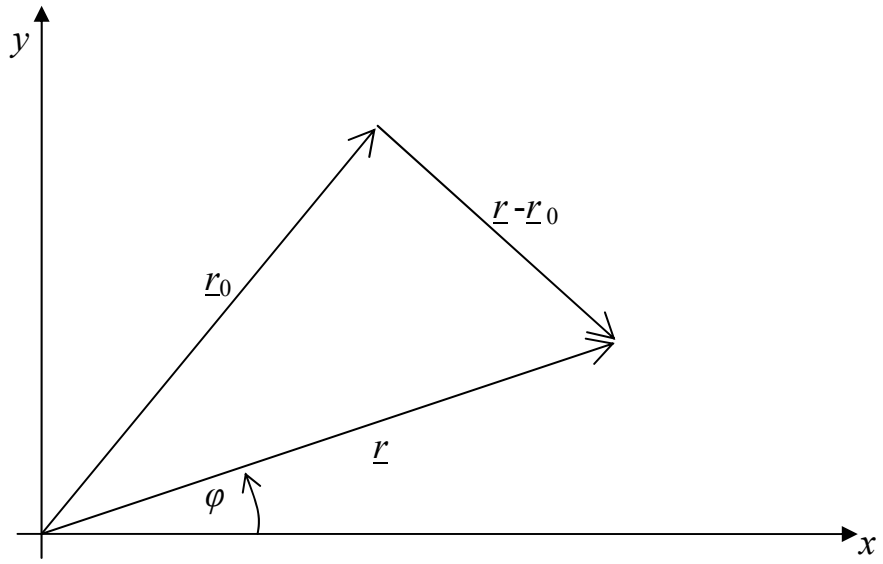


Figure 2.2 – Geometry.

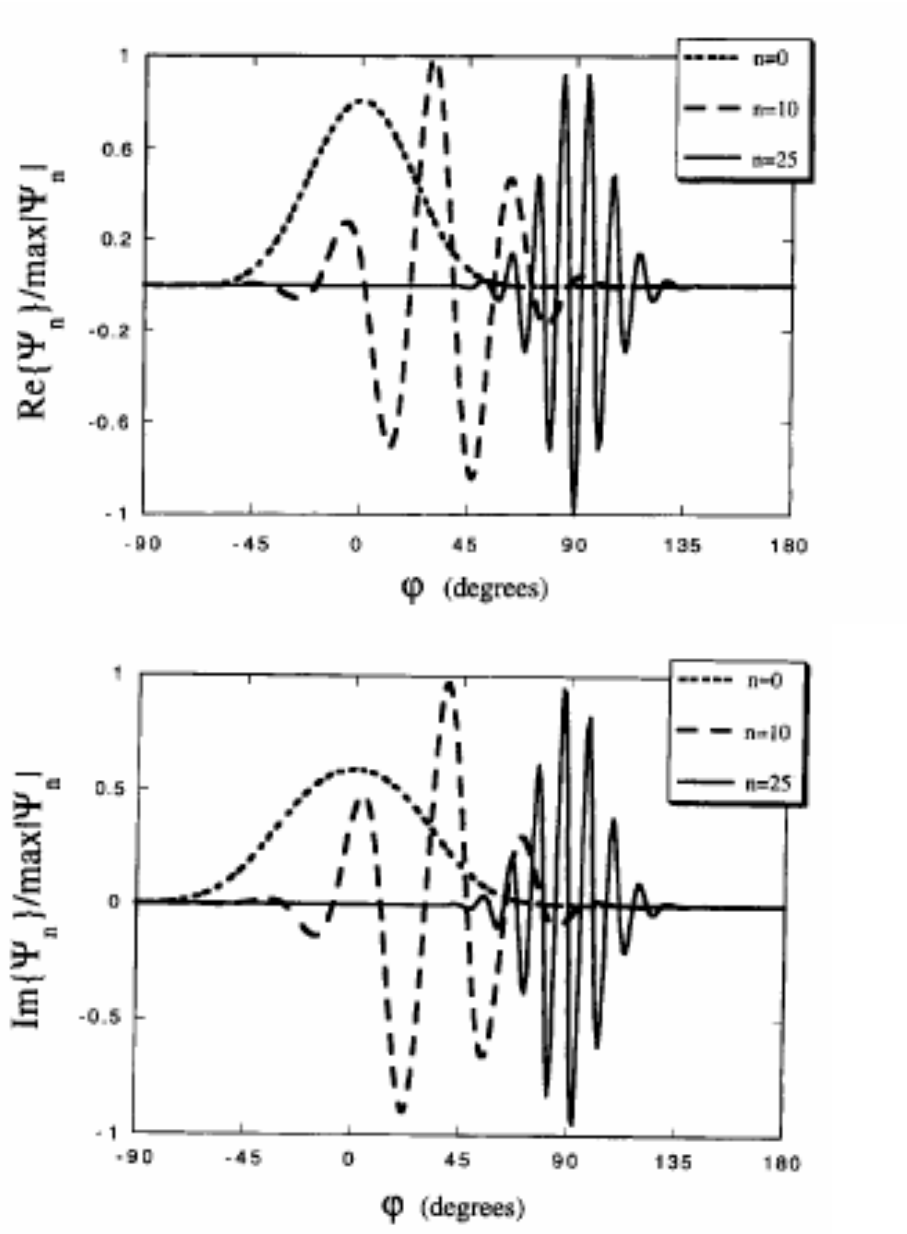
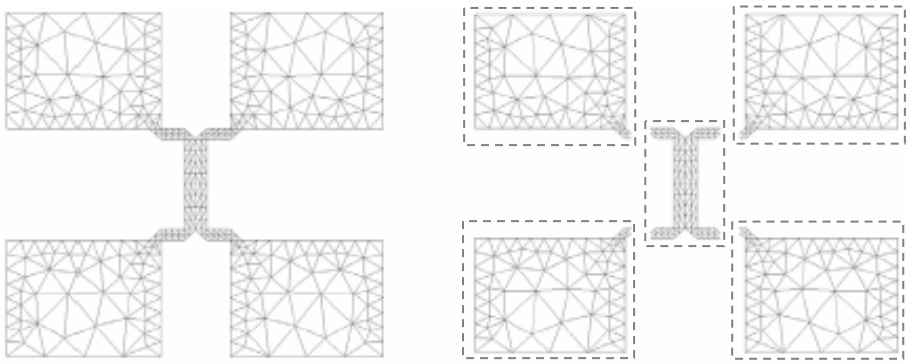


Figure 2.3 – Behaviour of the multipole fields $\Psi_n(\underline{r},\underline{r}_0)$*, for* $\underline{r}_0 = -j\hat{x}$ *and* $\underline{r} = 3\lambda\left(\hat{x}\cos\varphi + \hat{y}\sin\varphi\right)$

A qualitative analysis reveals that the multipoles produce well confined beams, propagating in the φ direction and with a spatial frequency that increases with the order n (see fig.2.3). Such behaviour resembles that of the Gabor expansion basis functions: substantially windowed exponentials.
The idea is to use a suitable set of multipoles that can constitute a complete base to express the field on the boundary. To do that, the positioning and the shift of the complex multipoles have to be choosen carefully to match the parameter set in the Gabor expansion for the window size, the shift along the boundary and the frequency sampling rate. This should be accomplished by choosing proper location and complex shifting for the complex multipoles, but cannot be made in a rigorous way, since just few heuristic guidelines are proposed. The advantage with respect to the direct use of the Gabor expansion is due to the fact that the simil-Gabors expansion functions are obtained analytically from the (2.4), thus avoiding massive numerical integrations [8]. Moreover, the complex multipoles radiated beams exhibits an unidirectional character, strongly reducing the coupling of each source with far parts of the boundary. This characteristic has two positive consequences: on the one hand, it provides a good sparsification of the matrix, with possible clipping of the lower elements and reduction of the computational costs; on the other hand, it is found that the condition number of the impedance matrix is reduced with respect to that usually obtained with isotropic or bidirectional sources.
According to the authors, by means of the CMBA it is possible to solve the problem of the scattering from a smooth surface with the use of a number of multipoles that approaches the Nyquist limit of 2 unknowns per wavelength.

2.3 The Synthetic Function Expansion

The Syntetic basis Function eXpansion (SFX), proposed by Matekovits and Vecchi [20,21,23], applies to the MoM-based analysis of array antennas, in particular to printed structures. The approach starts with a decomposition of the domain, including both the radiating and the beam forming network parts. The goal is to employ a few, "global" basis functions defined on different portion of the structure, able to represent the behaviour of any part even when it is embedded in the whole structure. The basic idea is to solve the problem in each "block" by supposing it isolated, but under the effect of a generic excitation.



a) *b)*

Figure 2.4 – Application example: 2x2 patch array; a) RWG mesh of the whole structure; b) domain decomposition

With reference to a patch array antenna, let us decompose the whole structure in a certain number of sub-domains (see fig. 2.4) [23]. The *n*_th block is studied by means of a standard MoM approach, exploiting F_n Rao-Wilton-Glisson basis functions, $\underline{f}_1$, $\underline{f}_2$, …, $\underline{f}_{F_n}$.

This leads to the evaluation of a complete F_n-by-F_n impedance matrix, $\underline{\underline{Z}}_n$, for the stand-alone block. Then, the isolated problem is solved by

considering two kind of excitations: the "natural" one, deriving from the patch feeding stripline, and the "coupling" ones, taking into account the possible excitation due to the radiation of other parts of the structure (see fig. 2.5). The problem then undergoes an SVD decomposition in order to identify and retain the K_n "independent" solutions ψ_1, ψ_2, …, ψ_{K_n}, each defined on the support of the n_th block and described by means of a summation of RWG functions:

$$\underline{\psi}_k = \sum_{\alpha=1}^{N_f} \psi_{k,\alpha} \underline{f}_\alpha \qquad (2.5)$$

The obtained synthetic functions ψ_1, ψ_2, …, ψ_{K_n} are evaluated for each block and constitute a set of basis functions for the analysis of the whole structure. In this way, we are "compressing" the original impedance matrix, by means of the expansion coefficients $\psi_{k,\alpha}$, provided that the number of SFs needed for each block is smaller than the number of required RWG functions (i.e. $K_n < N_{f,n}$) [22]. It is important to note that further savings in the computational effort arise from the fact that, when dealing with array antennas, many blocks are identical, thus the evaluation of the SFs can be done once for a large number of sub-domains.

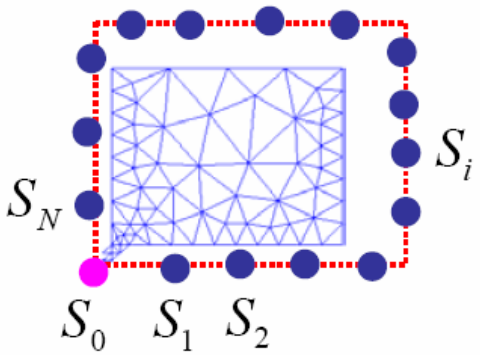


Figure 2.5 – "Natural" (S_0) and "coupling" (S_1,…, S_N) excitations considered on the single block

The technique has been applied to various cases of practical interest [20, 21, 22]. It has been shown that a satisfactory analysis of an array antenna can be carried out by means of a number of SF quite smaller than the original number of RWG basis functions

2.4 The Method of Auxiliary Sources

The Method of Auxiliary Sources was introduced and developed by several researchers in the University of Tbilisi, Georgia, and applied to the electromagnetic scattering issue, as well as to other kind of problems [14-19].
In typical integral-equation methods, like the Method of Moments, the problem of the scattering by a given object is formulated in terms of the known impressed field and unknown continuous currents distributed on the boundary surface. Conversely, the idea under the MAS is to directly reconstruct the scattered field on the boundary (and outside) by means of a set of simple point sources, called Auxiliary Sources, placed inside the volume scatterer. A popular choice is to use a set of elementary sources (filamentary currents in two-dimensional problems and elementary dipoles in three-dimensional ones) on a so-called Auxiliary Surface, placed just beneath the boundary surface: the fields radiated by such sources constitutes a non-orthogonal basis for the expression of the scattered field in the outer region (see figure 3.9). Analogously to the MoM, a linear system is built to relate the field radiated by the ASs with the incident field on a set of testing points placed on the boundary. The solution of the system, i.e. the inversion of the impedance matrix, provides the required values for the ASs excitations.
The MAS approach presents several advantages with respect to the standard MoM, in particular when dealing with large and smooth metallic scatterers. The fictitious sources adopted have a simple, analytically known expression, thus avoiding massive numerical integrations on the basis function domains required in classical MoM to evaluate the impedance matrix entries. Moreover, since the ASs do not lay on the boundary, there is no need to deal with the singularity of the Green's function, that elsewhere should be treated carefully.

Finally, due to reasons that will be clarified in the following chapters, the MAS analysis of a large metallic scatterer can be performed with a number of unknowns slightly larger than 2 per wavelength, while a classical MoM can require at least 10 unknowns per wavelength.
On the other hand, some drawbacks come out from the positioning of the ASs beneath the boundary. One is the increase of the impedance matrix condition number, and the consequent need for a good precision in numerical tools. But the main disadvantage of the MAS is probably the lack of rules for the optimal positioning of the ASs and the testing functions, that in literature are displaced following heuristic criteria.
This last issue will be approached in the following chapters, exploiting the theory on the optimal sampling of the scattered fields, thus providing some useful guidelines for the AS and testing point positioning. With these considerations, a MAS-based technique for the analysis of large scatterers will be implemented by exploiting a non-redundant number of unknowns [20].

2.5 The Impedance Matrix Localization Method

In 1990, Francis X. Canning proposed a novel approach to mitigate the well-known drawback of the MoM, the need for a large and dense impedance matrix and the consequent high storage and computational effort requirements. The key idea is to transform the MoM impedance matrix $\underline{\underline{Z}}$ by means of a suitable linear operator, thus effectively changing the basis functions and obtaining the sparsification of the matrix [21]. In fact, the linear system arising from a MoM-based method can be written as:

$$\underline{\underline{Z}}\,\underline{J} = \underline{E} \tag{2.6}$$

where $\underline{J}$ is the unknown vector of the induced current coefficients, and $\underline{E}$ is the excitation vector. In the IML method, a matrix $\underline{\underline{A}}$ is introduced, transforming the linear system (2.6) into:

$$\underline{\underline{A}}\underline{\underline{Z}}\left(\underline{\underline{A}}^T \underline{I}\right) = \underline{\underline{A}}\underline{E} \qquad (2.7)$$

that can be written as:

$$\underline{\underline{T}}\,\underline{I} = \underline{V} \qquad (2.8)$$

with $\underline{\underline{T}} = \underline{\underline{A}}\,\underline{\underline{Z}}\,\underline{\underline{A}}^T$, $\underline{J} = \underline{\underline{A}}^T \underline{I}$ and $\underline{V} = \underline{\underline{A}}\underline{E}$.
The goal is to find a proper transformation $\underline{\underline{A}}$ so that most of the elements of the matrix $\underline{\underline{T}}$ turns out to be almost zero, and can be cut under a certain threshold, resulting in a sparse matrix approximating the original problem.
To this purpose, the transformed matrix T is obtained by exploiting directional basis functions. With reference to a 2D problem, the boundary is conceptually broken into regions, and on each one of these *M* basis functions are defined as:

$$\exp(2\pi jhs/a) \qquad for\ h=0,\ldots,M\text{-}1 \qquad (2.9)$$

where *s* is the arc length and *a* the size of the region. The field radiated by a current distribution like (2.9) shows a maximum intensity for an direction θ (indicating the angle with the respect to the boundary normal) given by:

$$\sin(\theta) = (2\pi h/a)/k \qquad for\ h=0,\ldots,(M\text{-}1)/2 \qquad (2.10)$$

As a result of the directivity of the basis functions, a few interactions are very strong, and the correspondent elements in the impedance matrix have high magnitude, while most of the other elements in $\underline{\underline{T}}$ are almost zero.
This key idea has been improved with the add-on of a weighting operator, indicated as $\underline{\underline{W}}$, that transforms the expression of $\underline{\underline{T}}$ into:

$$\underline{\underline{T}} = \underline{\underline{A}}\left(\underline{\underline{W}}\,\underline{\underline{Z}}\,\underline{\underline{W}}^T\right)\underline{\underline{A}}^T \qquad (2.11)$$

In a first time, it was introduced an operator $\underline{\underline{W}}$ able to reduce the side lobes of the basis function radiation pattern, by actually modifying the current distributions (2.9) by means of a tapering function [21]. In a second time, the weighting concept has been generalized, and the matrix $\underline{\underline{W}}$ used to introduce other kinds of transformations, aiming to improve the performances of the method [23].
The numerical results of the application of the IML technique have shown that an N-by-N impedance matrix can be approximated by means of a sparse matrix having roughly $100N$ non-zero elements [21]. As a consequence, both the memory storage and the computational effort related to the evaluation and solution of the linear system are significantly reduced.

2.6 Wavelet-based approaches

An interesting branch of techniques effectively solving an integral formulated scattering problem involves the concept of wavelet representation, ideally following the Mallat's Multi-Resolution (MR) framework for the signal decomposition [35].
Two main schemes have been proposed, differing for the level of application of the MR approach. The first one solves the classical MoM problem by exploiting suitable wavelet-like basis functions, and goes by the name of Multi-Resolution MoM (MRMoM). The second one proposes the direct application of a wavelet transformation to the linear system arising from a standard MoM implementation, and is generally referred to as Discrete Wavelet Transform (DWT) method. In both cases, the goal is to obtain a sparse, well conditioned impedance matrix that can be stored and iteratively inverted with a reduced effort.

2.6.1 Multi-Resolution MoM

The idea behind the MRMoM approach is to use a wavelet-base set of basis function to solve those scattering problems that are formulated in the integral form [24]:

$$\int_L f(x')K(x,x')dx' = g(x), \qquad x \in L \qquad (2.12)$$

where K is the known kernel of the integral equation, f is the unknown response to be determined and g represents the known excitation. For the sake of simplicity, even if the (2.12) is in general vectorial and two- or tree-dimensional, in the following we will suppose it to be scalar, without loss of generality.
The MRMoM approach to the solution of the (2.12) is based on the use of a hierarchical basis function set. In accordance with the wavelet theory, where a *mother wavelet* is scaled and shifted to build a base for the domain in which the signals have to be represented, a basic function, $\psi(x)$, - is suitably chosen and replicated to build the basis functions, $\psi_n(x)$, that have to represent the solution of the integral equation.
The choice of the mother wavelet strongly influences the effectiveness of the method in the desired sparsification of the impedance matrix arising from the MoM solution of the equation (2.12). We say that the function $\psi(x)$ has a *vanishing moment of order m* if:

$$\int \psi(x)x^{m-1}dx \qquad (2.13)$$

In the hypothesis that the kernel K is smooth enough to be approximated by a polynomial expression of order M, if the basis functions $\psi_n(x)$ have vanishing moments of orders (1,…,M), most of the impedance matrix elements are expected to be very small [26]. Obviously, completely vanishing functions are unfit to represent a general problem. In practical implementations, also some non-vanishing "scaling" functions $\Phi_n(x)$ are introduced in the basis

function set, aiming to represent the spatial DC component of the solution. Moreover, some other functions with a reduced number of vanishing moments are used to represent the solution at the boundary limits, in addition to the previous ones that are appropriate for the inner boundary region.

In literature, many kinds of mother wavelets (and scaling functions) have been used. The simplest one is the Haar function, shown in figure 2.6 [25]. Other suitable kind of functions comes from the Chui (piecewise) wavelets (see figure 2.7) [26].

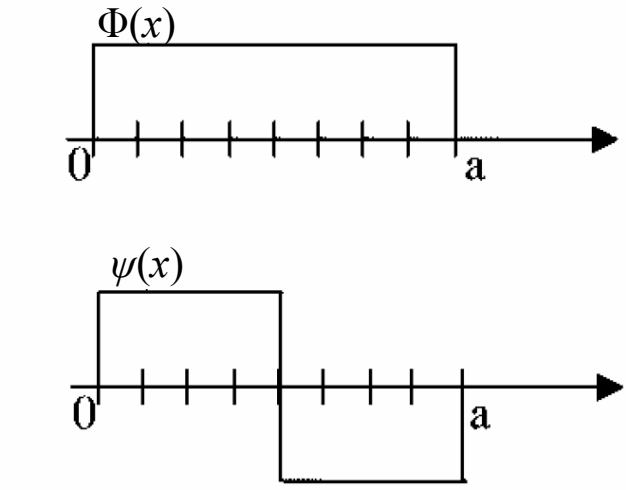


Figure 2.6 – The Haar wavelet function $\psi(x)$ *and the chosen scaling function* $\Phi(x)$

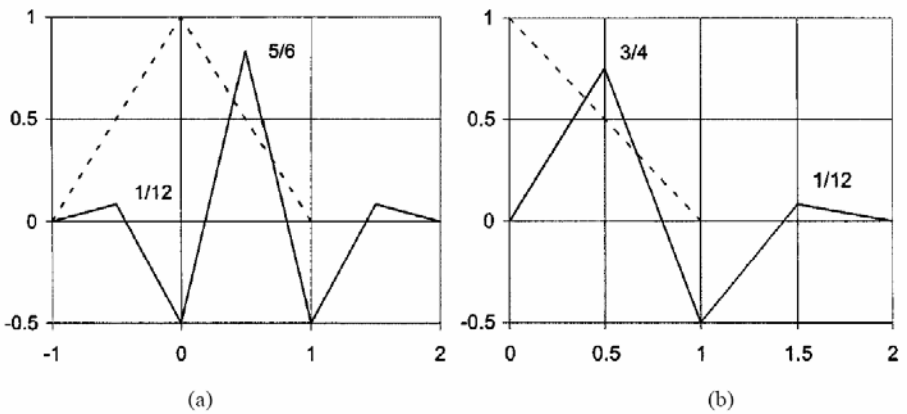


Figure 2.7 – Chui wavelet functions (solid lines) and scaling functions (dashed lines) in the inner (a) and left-boundary (b) case.

The chosen mother wavelet(s) are then dilated and replicated to build, together with the scaling function(s), a complete set of basis functions (see fig. 2.8, [25]). The higher is the hierarchy level, the smallest is the support of the functions included, and the finer is the representation provided.
It can be convenient for the reduction of the number of unknowns that, at higher levels, the basis functions do not recover the entire geometry, leaving uncovered the areas where the solution is expected to have lower spatial frequencies.

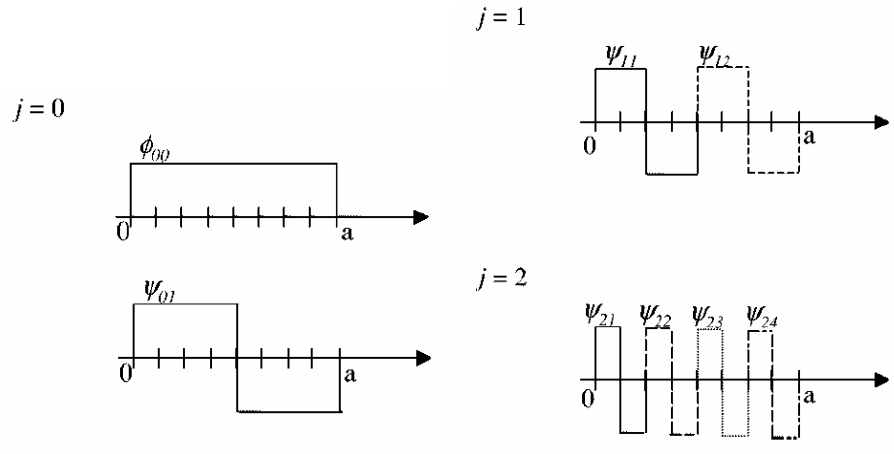
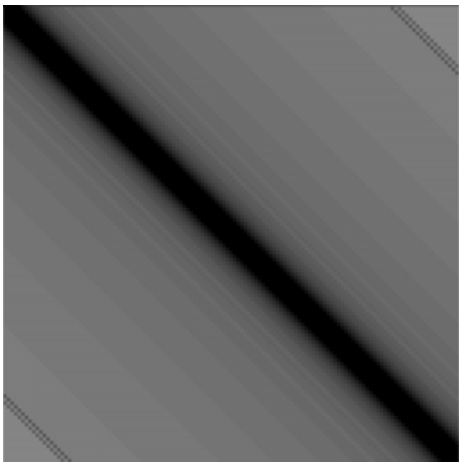
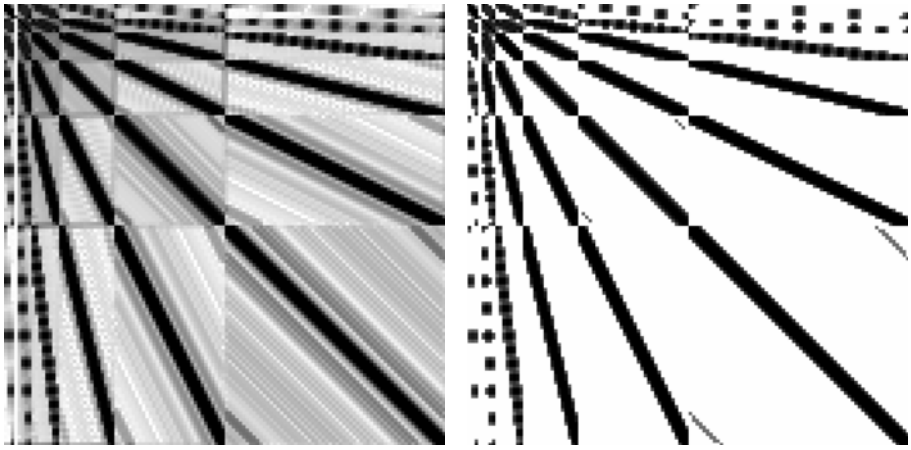


Figure 2.8 – A set of basis functions based on the Haar wavelet.

The adoption of the above described approach for the construction of the basis function set allows the MRMoM techniques to emphasize the dynamic of the impedance matrix entries, so that it is possible to clip the elements under a certain threshold, thus obtaining a sparse impedance matrix (see figure 2.9). As a result, the memory storage and the computational effort required to formulate and solve the integral equation (2.12) are considerably reduced.



a)



b) c)

Figure 2.9 – Example of impedance matrix sparsification: a) standard MoM impedance matrix; b) impedance matrix obtained by adopting a MRMoM approach; c) sparse matrix obtained by clipping the entries under a threshold

2.6.2 Discrete Wavelet Transformation approach

The Discrete Wavelet Transformation approach aims to solve a MoM formulated scattering problem by applying a suitable transformation to the related impedance matrix. Being $\underline{\underline{Z}}\underline{I}=\underline{E}$ the linear system to be solved, where $\underline{\underline{Z}}$ is the impedance matrix, $\underline{I}$ the unknown solution and $\underline{E}$ the known excitation, the problem is turned into the solution of:

$$\underline{\underline{Z}}'\,\underline{I}'=\underline{E}' \tag{2.14}$$

where

$$\underline{\underline{Z}}'=\underline{\underline{T}}\,\underline{\underline{Z}}\,\underline{\underline{T}}^{T}, \qquad \underline{E}'=\underline{\underline{T}}\,\underline{E}, \qquad \underline{I}'=\underline{\underline{T}}^{T}\,\underline{I} \tag{2.15}$$

where $\underline{\underline{T}}$ is an orthogonal matrix satisfying $\underline{\underline{T}}^{T}=\underline{\underline{T}}^{-1}$ [27].
The aim is to find a nonsingular matrix $\underline{\underline{T}}$ so that the transformed impedance matrix $\underline{\underline{Z}}'$ contains many small elements that can be neglected without largely affecting the solution $\underline{I}'$ [29]. In this way, the impedance matrix becomes sparse, and can be stored and inverted with reduced effort. In some sense, we are changing the base of the Z operator: the idea is similar to the one outlined in the previous section, about the MRMoM, but the transformation is applied directly on the impedance matrix rather than on the basis function set.
The matrix $\underline{\underline{T}}$ is built as basis changing operator, and its rows are wavelet vectors representing a possible basis of R^{N} [27]. As *mother wavelet*, the Debauchies wavelet has often been used (see fig. 2.10). An example of transformation matrix obtained by means of this function is shown in figure 2.11 [27], where the different bands in the image correspond to the different resolution scales in the basis. The sparsification effectiveness is shown in figure 2.12 where a transformed matrix $\underline{\underline{Z}}'$ is compared to the dense standard MoM matrix for the same problem [27].
The numerical analysis has shown that a properly chosen wavelet transformation provides a sparse impedance matrix containing roughly $N\log N$ non-zero elements, thus reducing the operation number of the

matrix-by-vector multiplications involved in the iterative solver to O(*N*log*N*) [7].

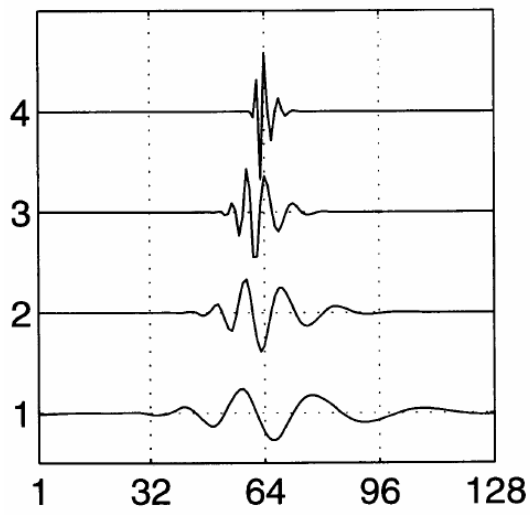


Figure 2.10 – Debauchies wavelets

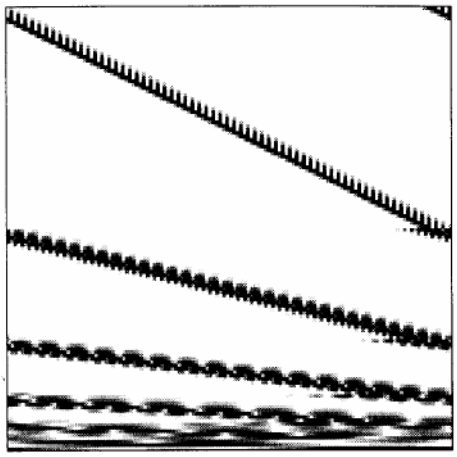
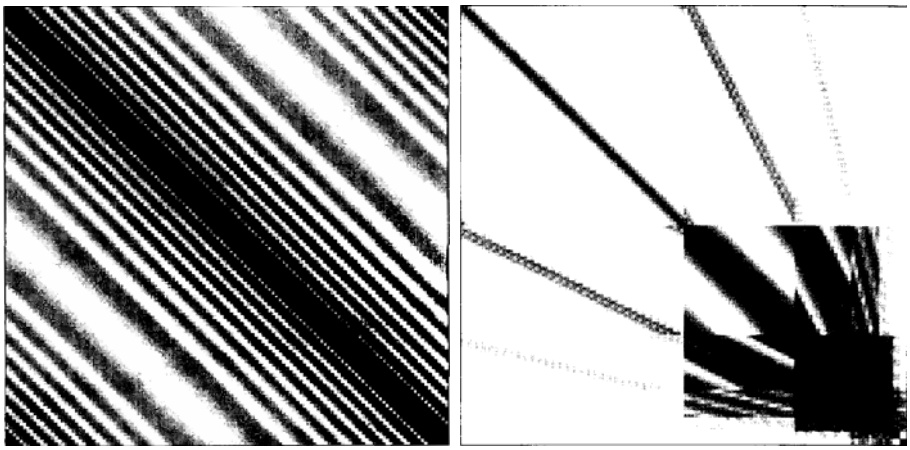


Figure 2.11 – Transformation matrix obtained by exploiting Debauchies wavelets



a) *b)*

Figure 2.12 – a) example of dense standard MoM matrix; b) impedance matrix obtained by wavelet-based transformation

2.7 The CG-FFT approach

As said before, the bottleneck of the iterative processes used to solve the linear system arising from the MoM formulation of the scattering problem is the matrix by vector multiplication, where the impedance matrix is dense. Such multiplication represents the evaluation of the field radiated by the currents induced on the scatterer, and can be written as a convolution integral:

$$\int_B g(\underline{r}-\underline{r}')\, j(\underline{r}')\, d\underline{r}' \qquad (2.16)$$

where $g(\underline{r})$ is the Green's function, $j(\underline{r}')$ is the induced current and B is the domain where the induced currents lay. Since the Green's function is translationally invariant, the evaluation of the convolution (2.16) can be expedited by using an FFT scheme.

The approach is particularly attractive for the analysis of linearly extended structures like conducting strips, and periodic structures like antenna arrays. In fact, when identical parts of the geometry reside on a regular grid, the pertinent MoM impedance matrix has a block-Toeplitz structure. Exploiting the Toeplitz structure, it is possible to perform the matrix-by-vector multiplication in O(Nlog$N$) operations by FFT [30].
For the sake of simplicity, let us consider an uni-dimensional problem concerning the current distribution on a wire, approached by means of the canonical MoM expansion in triangular basis functions (see figure 2.13).

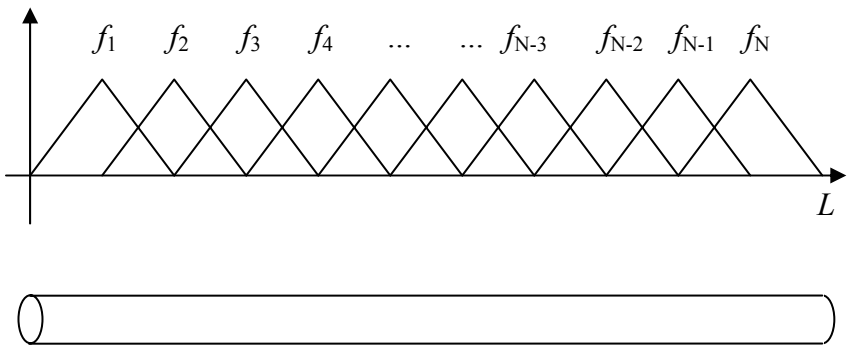


Figure 2.13 – The wire domain discretized by means of triangular basis functions

It is clear that the $z_{n,m}$, element of the impedance matrix, representing the interaction between the n_th basis and the m_th test function, just depends on the difference (n-m), and can be indicated as z_{n-m}. In this way, the impedance matrix just contains 2N-1 independent entries (see fig. 2.14).

$$\underline{\underline{Z}} = \begin{bmatrix} z_0 & z_1 & z_2 & z_3 & \dots & z_{3-N} & z_{2-N} & z_{1-N} \\ z_{-1} & z_0 & z_1 & z_2 & z_3 & \dots & z_{3-N} & z_{2-N} \\ z_{-2} & z_{-1} & z_0 & z_1 & z_2 & z_3 & \dots & z_{3-N} \\ z_{-3} & z_{-2} & z_{-1} & z_0 & z_1 & z_2 & z_3 & \dots \\ \dots & z_{-3} & z_{-2} & z_{-1} & z_0 & z_1 & z_2 & z_3 \\ z_{N-3} & \dots & z_{-3} & z_{-2} & z_{-1} & z_0 & z_1 & z_2 \\ z_{N-2} & z_{N-3} & \dots & z_{-3} & z_{-2} & z_{-1} & z_0 & z_1 \\ z_{N-1} & z_{N-2} & z_{N-3} & \dots & z_{-3} & z_{-2} & z_{-1} & z_0 \end{bmatrix}$$

Figure 2.14 – Toeplitz structure of the impedance matrix

The first observation concerns the possibility of storing just $2N$-1 values instead of N^2. But another advantage arises in the evaluation of the matrix by vector multiplication, as clarified in the following.
In fact, a new ($2N$-1)-by-($2N$-1) matrix $\underline{\underline{C}}$ can be built by adding N-1 rows and columns to $\underline{\underline{Z}}$:

$$\underline{\underline{C}} = \begin{bmatrix} \underline{\underline{Z}} & \underline{\underline{A}} \\ \underline{\underline{B}} & \underline{\underline{C}} \end{bmatrix} \tag{2.17}$$

where the joined sub-matrices are called $\underline{\underline{A}}$, $\underline{\underline{B}}$ and $\underline{\underline{D}}$.

It is easy to build $\underline{\underline{C}}$ as a *circulant* matrix by properly choosing the elements of $\underline{\underline{A}}$, $\underline{\underline{B}}$ and $\underline{\underline{D}}$ among the (2*N*-1) independent entries in $\underline{\underline{Z}}$, as depicted in figure 2.15. The obtained circulant matrix is defined by the only (2*N*-1) elements of the first column $\underline{c}$, since all the other columns can be derived from it by means of a simple shift.

$$\left[\begin{array}{ccccc|cccc}
z_0 & z_1 & \dots & z_{2-N} & z_{1-N} & z_{N-1} & z_{N-2} & \dots & z_{-1} \\
z_{-1} & z_0 & z_1 & \dots & z_{2-N} & z_{1-N} & z_{N-1} & z_{N-2} & \dots \\
\dots & z_{-1} & z_0 & z_1 & \dots & z_{2-N} & z_{1-N} & z_{N-1} & z_{N-2} \\
z_{N-2} & \dots & z_{-1} & z_0 & z_1 & \dots & z_{2-N} & z_{1-N} & z_{N-1} \\
z_{N-1} & z_{N-2} & \dots & z_{-1} & z_0 & z_1 & \dots & z_{2-N} & z_{1-N} \\
\hline
z_{1-N} & z_{N-1} & z_{N-2} & \dots & z_{-1} & z_0 & z_1 & \dots & z_{2-N} \\
z_{2-N} & z_{1-N} & z_{N-1} & z_{N-2} & \dots & z_{-1} & z_0 & z_1 & \dots \\
\dots & z_{2-N} & z_{1-N} & z_{N-1} & z_{N-2} & \dots & z_{-1} & z_0 & z_1 \\
z_1 & \dots & z_{2-N} & z_{1-N} & z_{N-1} & z_{N-2} & \dots & z_{-1} & z_0
\end{array}\right]$$

Figure 2.15 – Structure of the circulant matrix $\underline{\underline{C}}$

Our aim is to evaluate the product $\underline{\underline{Z}}\,\underline{I} = \underline{E}$. If the matrix $\underline{\underline{C}}$ is multiplied by a vector $\underline{V} = \begin{bmatrix} \underline{I} \\ \underline{0} \end{bmatrix}$, where $\underline{0}$ is a null vector of length *N*-1, we obtain:

$$\underline{\underline{CV}} = \begin{bmatrix} \underline{\underline{Z}} & \underline{\underline{A}} \\ \underline{\underline{B}} & \underline{\underline{C}} \end{bmatrix} \cdot \begin{bmatrix} \underline{I} \\ \underline{0} \end{bmatrix} = \begin{bmatrix} \underline{\underline{Z}}\,\underline{I} \\ \underline{\underline{B}}\,\underline{I} \end{bmatrix} = \begin{bmatrix} \underline{E} \\ \underline{\underline{B}}\,\underline{I} \end{bmatrix} \tag{2.18}$$

Due to the circulant structure of the matrix $\underline{\underline{C}}$, it turns out that the product 2.18 (and, as a consequence, the solution vector $\underline{E}$) can be evaluated as:

$$\begin{bmatrix} \underline{E} \\ \underline{\underline{B}}\,\underline{I} \end{bmatrix} = \underline{\underline{CV}} = DFT^{-1}\left[DFT\{\underline{c}\} \circ DFT\begin{bmatrix} \underline{I} \\ \underline{0} \end{bmatrix} \right] \tag{2.19}$$

The Discrete Fourier Transforms can be performed by means of an FFT algorithm (if necessary, by suitably adding padding elements in the $\underline{c}$ and $\underline{I}$ vectors to reach a length equal to 2^M) requiring a computational complexity of $O(N\log_2 N)$.
The procedure outlined for the above 1D case can be extended to other kind of geometries [30, 31, 32], also in 2D and 3D, with similar results in term of memory saving and reduction of the computational complexity.

2.8 The Fast Multipole Method

The Fast Multipole Method (FMM), proposed by Greengard and Rokhlin [33, 36, 37, 38], allows to reduce the memory storage and to speed up the matrix-by-vector multiplication involved in the iterative solution of the MoM linear system, for arbitrary scattering geometries.
The MoM impedance matrix is in general dense, since its entries represent all the N^2 possible interactions between any of the N basis and any of the N testing functions. As a consequence, the product of this matrix by a tentative-solution vector requires N^2 operations.
The FMM approach moves from the consideration that the MoM impedance matrix contains highly redundant information: this is confirmed by the fact that many block inside it, representing the interactions between far away parts of the geometry, are practically

low-rank matrices, and can be represented in an aggregated way. To this purpose, the scattering geometry is subdivided into small regions, each containing M nodes representing the localized basis (and testing) functions. In this way, the N-by-N impedance matrix $\underline{\underline{Z}}$ can be split into the sum of two matrices:

$$\underline{\underline{Z}} = \underline{\underline{Z}}' + \underline{\underline{Z}}'' \tag{2.20}$$

$\underline{\underline{Z}}'$ taking into account the interaction between basis and testing belonging to adjacent groups, $\underline{\underline{Z}}''$ containing the blocks that express the interaction between non adjacent clusters (see figure 4.3).
The elements of $\underline{\underline{Z}}'$ are evaluated like the entries of the standard MoM impedance matrix. If the number M of cluster is assumed approximately equal to the square root of N, the $\underline{\underline{Z}}'$ matrix only contains a number of non-zero elements proportional to $N^{1.5}$.
The matrix $\underline{\underline{Z}}''$ is not computed explicitly, but a reduced-redundancy representation is proposed for its elements. In fact, let us consider a generic element of $\underline{\underline{Z}}''$, $z_{n,n'}$, representing the interaction between the n'_th basis, belonging to the m'_th group, and the n_th testing function, belonging to the m_th group. It can be seen that a proper factorization of the Green's function leads to the following equivalence:

$$z_{n,n'} \approx \sum_{r=1}^{R} v_{r,n',m'}\, t_{r,m,m'}\, u_{r,n,m} \tag{2.21}$$

where:
- $u_{r,n,m}$ depends only on the distance between the n_th testing function and the center of the m_th cluster;
- $v_{r,n',m'}$ depends only on the distance between the n'_th basis function and the center of the m'_th cluster;
- $t_{r,m,m'}$ depends only on the distance between the centers of the m_th and m'_th (non-adjacent) clusters.
- R is a constant, proportional to $\sqrt{N}$ if $M \approx \sqrt{N}$.

It is easy to find out that the memory occupation of all the *translation functions* $t_{r,m,m'}$, all the *aggregation vectors* $v_{r,n,m}$ and all the *disaggregation vectors* $u_{r,n',m'}$, able to represent any non-zero element $z_{n,n'}$ of the matrix $\underline{\underline{Z}}''$, is proportional to $N^{1.5}$.
Once addressed, the problem of reducing the storage requirements of the impedance matrix, we want to speed-up the matrix by vector multiplication involved in the iterative solution algorithm, whose direct evaluation has a computational complexity of $O(N^2)$. In FMM, we have to compute:

$$\underline{\underline{Z}}\,\underline{x} \approx \underline{\underline{Z}}'\,\underline{x} + \underline{\underline{Z}}''\,\underline{x} \qquad (2.22)$$

The product $\underline{\underline{Z}}'\,\underline{x}$ is classically carried out and involves an amount of operation equal to the number of non zero elements of the sparse matrix $\underline{\underline{Z}}'$, i.e. proportional to $N^{1.5}$. On the other hand, the term $\underline{\underline{Z}}''\,\underline{x}$, by exploiting the representation 2.21, is evaluated in three steps:

- for each cluster *m*', the *aggregation* of the contributes due to the *N/M* basis function it contains;
- for each couple of basis and testing cluster, the evaluation of $t_{r,m,m'}$, representing the aggregated interactions;
- for each cluster *m*, the *disaggregation* of the contributes into the *N/M* testing function it contains.

All the three steps involve a number of operations proportional to $N^{1.5}$.
In some sense, we have turned the classical matrix-by-vector product, that can be seen as "full connected" network between *N* nodes, into a scheme featuring a level of suitable "hubs" (the cluster centers), thus reducing the number of links required (see figure 4.5).
In conclusion, the FMM approach allows to reduce the storage requirements of a MoM-like technique from N^2 to $N^{1.5}$, and the computational complexity of the matrix-by-vector product from $O(N^2)$ to $O(N^{1.5})$. Further improvements involve the use of more than one level of "hubs", thus achieving a memory requirement and a computational complexity that tends to $N\log N$ and $O(N\log N)$, respectively [39, 40].

For its flexibility and the possibility to be applied to various Integral-Formulation based technique for the scattering evaluation, the FMM has been chosen to speed up our non-redundant implementation of the Method of Auxiliary Sources.

Chapter 3

The Method of Auxiliary Sources

The fundamentals of the Method of Auxiliary sources are due to the Georgian mathematicians V. Kupradze, M. Aleksidze and I. Vekua [14, 15, 16]. Moreover, the technique was developed by D. Karkashadze, R. Zaridze and other researchers from the University of Tbilisi [17, 18, 19]. The first versions of the MAS were born with different names: "Method of Generalized Fourier Series", by Kupradze [14], "Method of Expansion in Terms of Metaharmonic Functions", by Vekua [15], "Method of Expansion by Fundamental Solutions", by Aleksidze [16]. Analogous results were independently found and proposed by Leviatan and Boag under the name of Generalized Formulation [41, 42].

The method of auxiliary sources has been successfully applied to a variety of problems: acoustics, hydrodynamics, electromagnetics [19]. We will focus our attention on the latter of these topics, in particular on the problem of the evaluation of the electromagnetic field scattered by a large object.

While in typical integral equation methods, like the Method of Moments, the boundary-value problem is solved in terms of continuous equivalent currents on the boundary surface, the idea under the MAS is to directly reconstruct the field on the boundary by means of a set of simple point sources, called Auxiliary Sources.

Our aim is to evaluate the field scattered by a metallic structure in air: this situation can be approximated with a Perfect Electric Conducting (PEC) object in free space. In the following, typical application of the MoM and the MAS techniques to this problem are presented, and some advantaged and disadvantages are discussed.

3.1 The Method of Moments

Let us consider a PEC object immersed in free space and illuminated by and incident field $\underline{E}_i$ (see figure 3.1). In the following, the harmonic *exp(jωt)* time dependence is assumed and suppressed. Let *S* be the (limited and closed) surface of the object, $\hat{n}$ the outward unit vector normal to the boundary, and $\underline{E}_1$, $\underline{H}_1$ and $\underline{E}_2$, $\underline{H}_2$ the electric and magnetic field outside and inside *S*, respectively. Since the internal medium is a perfect conductor, the field inside is equal to zero, and the boundary condition are written as:

$$\hat{n} \times \underline{E}_1 = 0, \quad \hat{n} \times \underline{H}_1 = \underline{J}_S$$
$$\hat{n} \cdot \underline{D}_1 = \rho_S, \quad \hat{n} \cdot \underline{B}_1 = 0 \qquad (3.1)$$

The resolution of scattering problem essentially consist in determining the surface sources (currents and charges) that are excited on the boundary by the known incident field $\underline{E}_i$. Once determined these current, the scattered field $\underline{E}_s$ can be easily evaluated anywhere.

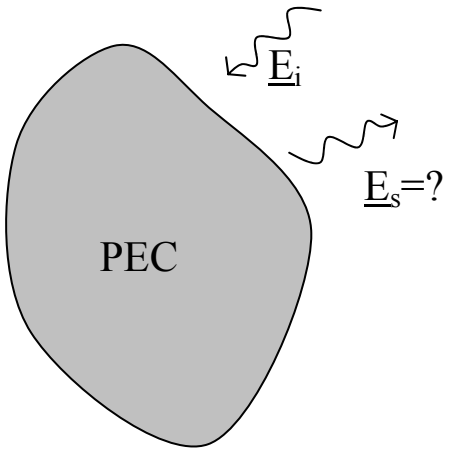


Figure 3.1. – PEC object scattering problem

Equivalently, the Maxwell's equations with the boundary conditions (3.1) must be solved. In the techniques usually referred as "MoM techniques", the differential equations are solved with an integral approach: the known incident field is related to the unknown sources by means of an integral relation, and this is solved by means of a mathematical method, namely, the Method of Moments [43].
It must be stressed that although the name "Method of Moments", should indicates just the mathematical technique able to solve a system of linear equations $L(f)=g$ (by using suitable expansion basis for the unknown function f), it is commonly used also to name the whole integral-approach scattering evaluation techniques. As an example, the Finite Element Method (FEM), an approach of different kind with respect the IE one, also relies on a form of MoM for the solution of the involved system of linear equations .

3.1.1 Integral Equation Formulation

The electromagnetic field can be expressed in terms of electric scalar potential V and magnetic vector potential $\underline{A}$. In the phasor domain, we have:

$$\underline{E} = -j\omega\underline{A} - \nabla V \quad , \tag{3.2a}$$

$$\mu\underline{H} = \nabla \times \underline{A} \tag{3.2b}$$

with the Lorentz gauge $\nabla \cdot \underline{A} = -j\omega V$.
These potentials are related to the surface field sources by the integral equations (see figure 3.2):

$$\underline{A}(r) = \mu \int_S \underline{J}_S(\underline{r}')g(\underline{r},\underline{r}')dr' \tag{3.3}$$

$$V(r) = \frac{1}{\varepsilon}\int_S \rho_S(\underline{r}')g(\underline{r},\underline{r}')dr' \tag{3.4}$$

where g is the Green's function, solution of the inhomogeneous Helmholtz equation, i.e.:

$$g(\underline{r},\underline{r}') = \frac{\exp(-jk|\underline{r}-\underline{r}'|)}{4\pi|\underline{r}-\underline{r}'|} \tag{3.5}$$

with $k = j\omega\sqrt{\varepsilon\mu}$.

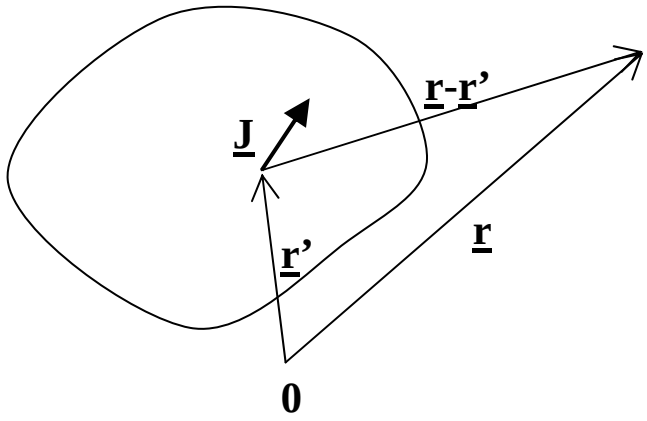


Figure 3.2 – Coordinate system for the evaluation of potentials

In the case of 2D geometries, sources are intended to be infinite filamentary currents and charge distribution, for which the Green's function reduces to:

$$g(\underline{r},\underline{r}') = -\frac{j}{4}H_0^{(2)}(k|\underline{r}-\underline{r}'|) \tag{3.6}$$

$H_0^{(2)}(\cdot)$ being the Hankel function of second kind and order zero [44].

From the (3.2a), we can express the electric field in terms of the only vector potential:

$$\underline{E} = -j\omega\left(\underline{A} + \frac{1}{k^2}\nabla\nabla\cdot\underline{A}\right) \tag{3.7}$$

Let's come back to the scattering PEC object immersed in free space and illuminated by an incident electric field $\underline{E}_i$. The boundary condition ensures the continuity of the tangential component of the electric field on the interface between medium. Since the total electric field on the boundary is the sum of the incident field $\underline{E}_i$ and the scattered field $\underline{E}_s$ (due to the sources on S), the tangential scattered electric field is equal and opposite to the incident field, i.e.,

$$n\times\left(\underline{E}_i\right)_S + n\times\left(\underline{E}_s\right)_S = 0 \tag{3.8}$$

By expressing the electric field in terms of the currents J_s by means of eq. (3.7) and by enforcing the condition (3.8), we are leaded to the formulation of a linear problem $L(f)=g$, where g is related to the known incident electric field, the function f stands for the unknown current distribution and L represents the radiation operator.
In particular, for points $P(\underline{r})$ on S, we have:

$$\hat{n}\times\mu\int_S j\omega(\underline{\underline{I}} + \frac{\nabla\nabla}{k^2})\cdot\underline{J}_S(\underline{r}')g(\underline{r},\underline{r}')dr' = \hat{n}\times\underline{E}_i(\underline{r}) \tag{3.9}$$

Since the expression (3.9) is obtained in terms of the electric field, by means of the expression (7), this is called Electric Field Integral Equation (EFIE) formulation.
Analogously, by applying the boundary condition (3.1) on the magnetic field instead of the electric field, the Magnetic Field Integral Equation (MFIE), since the total magnetic field is the sum of the incident field $\underline{H}_i$ and the scattered field $\underline{H}_s$, we can write the condition:

$$n \times \left(\underline{H}_i \right)_S + n \times \left(\underline{H}_s \right)_S = \underline{J}_S \qquad (3.10)$$

Exploiting equations (3.3), (3.2b) and we obtain the final expression for the MFIE formulation:

$$\underline{J}_S(\underline{r}) - \hat{n} \times \int_S \nabla \times \underline{J}_S(\underline{r}') g(\underline{r}, \underline{r}') dr' = \hat{n} \times \underline{H}_i(\underline{r}), \qquad (3.11)$$

with $\underline{r} \in S'$. An alternative to the EFIE ad MFIE formulation is the Combined Field Integral Equation. It is simply a linear combination of the equations (3.9) and (3.11):

CFIE = α EFIE + (1 - α) MFIE

Where α is a scalar value between 0 and 1. The utility of the CFIE formulation arise from the fact that both the EFIE and the MFIE suffer from internal resonances that lead to wrong results at specific frequencies. In this way, two boundary conditions are applied at the same time and the null-space solution of the resulting equation is zero for all frequencies. Another advantage of the CFIE is that it usually provides better conditioned matrices.

3.1.2 MoM solution of the linear system

The Moment Method is now described. We are concerned with a linear operator equation in the form

$$L(f)=g \qquad (3.12)$$

L being and integral-differential operator, *g* being a known function, *f* unknown: the MoM will provide an effective approach for the numerical solution of this problem.

The basic idea of the MoM is to expand the unknown function f into a set of linear independent basis functions f_n, suitably chosen, so that, retaining only N terms of the expansion, we can approximate:

$$f \approx \sum_{n=1}^{N} \alpha_n f_n \tag{3.13}$$

α_n being unknown coefficients to be determined. The linearity of the operator L ensures that

$$L\left(\sum_{n=1}^{N} \alpha_n f_n \right) = \sum_{n=1}^{N} \alpha_n L(f_n) \tag{3.14}$$

$$\sum_{n=1}^{N} \alpha_n L(f_n) \approx g \tag{3.15}$$

In general, due to the finite number of terms in (3.13), the above expression does not allows an exact solution. Anyway, a set of values for the coefficients α_n can be find matching at best the equivalence (3.15).
This is obtained by projecting both terms of the equation into a space of N weighting functions (w_1 , ... w_N) . The equation (3.12) can be rewritten in the so called "weak form", because the equivalence is enforced only on the subspace spanned by the N weighting functions w_m, $m=1,...,N$, i.e.,

$$\sum_{n=1}^{N} \alpha_n < w_m, L(f_n) >=< w_m, g > \tag{3.16}$$

where $< \cdot,\cdot >$ indicates the scalar product.
Accordingly, a linear system of N equation in N unknowns is obtained.
Usually, the equation (3.13) is rewritten in the form:

$$\underline{\underline{Z}}\,\underline{\alpha} = \underline{b} \tag{3.17}$$

where in $\underline{\alpha}$ is the vector of the unknown coefficients, the matrix $\underline{\underline{Z}}$, the *Impedance Matrix* is, is made by the elements $z_{m,n}$ given by :

$$z_{m,n} = < w_m, L(f_n) > \tag{3.18}$$

and the vector $\underline{b}$ is given by

$$b_m = < w_m, g > \tag{3.19}$$

The solution of the linear system (3.17) is equivalent to the inversion of the impedance matrix $\underline{\underline{Z}}$.

3.1.3 Properties of the impedance matrix

The MoM impedance matrix $\underline{\underline{Z}}$ is a square, dense matrix accounting for the interactions between any current expansion function and any testing function on the scattering boundary. The straightforward solution of the system (3.17), involving the inversion of the matrix Z, usually demands an high computational effort. Being NxN the size of Z, the exact solution is obtained by means of algorithms such the LU decomposition or the Gaussian elimination, all of them requiring a computational effort equal to $O(N^3)$.
As an alternative, in order to speed up the inversion process, iterative methods providing an approximate solution are frequently adopted. Such algorithms build a sequence of tentative solutions, that are improved at each step, until the required precision is achieved. Since any iteration is essentially constituted by a matrix-by-vector product, assuming that the number of steps needed to attain the convergence is essentially independent on the problem size, the computational complexity of this inversion process is of the order, $O(N^2)$.
Moreover, taking into account the physical meaning of the matrix-by-vector product, i.e., the evaluation of the field radiated by the current distribution on the boundary, the researcher were leaded to the

development of techniques able to reduce the relative computational complexity from $O(N^2)$ to $O(N^{3/2})$ and up to $O(N\log N)$ [7]. Among these, the Fast Multiple Method has been considered and applied in the next chapters.

3.1.4 Basis and testing functions

The choice of a suitable set of basis and testing function is an important issue for the MoM-based techniques. In fact, to improve the efficiency of the method, the number of basis and testing function should be kept as small as possible. This means that a relatively small number of expansion function should be able to satisfactory approximate the solution of the physical problem. Analogously, a small number of testing functions should be able to approximate the equivalence of the left and right hand of the equation (3.16). Often, the testing function are chosen coincident with the basis functions: thus obtaining the Galerkin's method. Other very simple testing functions used are the Dirac's delta function. With this choice the boundary conditions are enforced on a discrete set of points (point-matching): this is also called "collocation method".
With reference to the domain D of f (in our case, the surface S), two types of basis and testing functions can be adopted: the sub-domain and the entire-domain functions.
In the first case, the domain is subdivided into a large number of sub-domains, wherein the each basis function is defined. This approach allows to easily handle any geometry, which can be easily decomposed in small sub-domains.
Entire-domain basis and testing functions are instead defined on the whole D: for example, they have a sinusoidal or polynomial behaviour. Sometimes the domain D is subdivided into several smaller domains (but always larger than the ones found in the sub-domain approach), and entire-domain functions are defined on each one. In antennas, this is usually done when the object can be decomposed in several elementary parts for which characteristic expansion functions are known.

3.1.5 Examples of basis functions

The continuous current distribution on the boundary has to be approximated by a weighted sum of a finite number of basis functions.
Usually, first the domain itself must be suitably approximated. As an example, in the 2D geometry, the scatterer line D is approximated by a set of segments, say $D_1,\ldots,D_N$ (see fig 3.3), on which the basis functions will be defined.
The Dirac's functions, the simplest basis functions, are rarely chosen both for the poor accuracy they provide and for their singular behaviour, making unfeasible any derivative(see eq. (3.9)). Anyway, as said before, delta functions are often used as testing functions: this means that the equation (3.8) is enforced by means of a point-matching technique (collocation method).
Other simple basis functions are the pulse functions, defined as:

$$f_n = \begin{cases} 1 & in\ the\ n\_th\ subdomain \\ 0 & elsewhere \end{cases} \qquad (3.20)$$

The use of pulse functions provides a staircase discontinuous approximation of the current distribution (see fig. 3.4).
Among the continuous basis functions, the simplest ones are the so called "rooftop" basis functions providing a piece wise approximation of f. These functions are defined on two adjacent subdomains, are of triangular shape , are equal to 1 on the edge between subdomains and decreases linearly down to 0 at the external boundaries of the subdomains (see fig. 3.5).
If needed, finer approximations are possible, for instance by using basis functions with more degrees of regularity.
In the 3D geometry, the boundary surface can be subdivided and approximated by a set of subdomains, usually triangular patches.
As in the 2D case, it is simple to obtain a linear interpolation of the current distribution. Once the domain D has been approximated as a sum of triangles (fig. 3.6), the support of the n_th basis function is

given by the set of triangles that gather around the n_th node, being equal to 1 on the central node and decreasing linearly to 0 at the external boundary of each triangle (see fig. 3.7)
It is worth noting that the most used basis functions are the Rao-Wilton-Glisson (RWG) edge elements (fig. 3.8) [34]. The boundary surface is approximated by a sum of triangles, and on each pair of triangles (i.e. for each edge) the basis function is defined as:

$$f(\underline{r}) = \begin{cases} (l/2A^{+})\underline{\rho}^{+}(\underline{r}) & \underline{r} \in T^{+} \\ (l/2A^{-})\underline{\rho}^{-}(\underline{r}) & \underline{r} \in T^{-} \\ 0 & elsewhere \end{cases} \qquad (3.21)$$

where T^{+} and T^{-} indicates the two triangles, l the edge length, $\underline{\rho}^{+}$ the vector connecting the free vertex of T^{+} with the observation point $\underline{r}$, $\underline{\rho}^{-}$ the vector connecting the free vertex of T^{-} with the observation point $\underline{r}$, and A^{+} and A^{-} the areas of T^{+} and T^{-}, respectively. The edge between T^{+} and T^{-} univocally defines the couple of triangles. With the shown choice for $f(\underline{r})$, the current is continuous inside each triangle, and purely parallel along the non-defining edges. Moreover, the current normal to the defining edge is continuous.

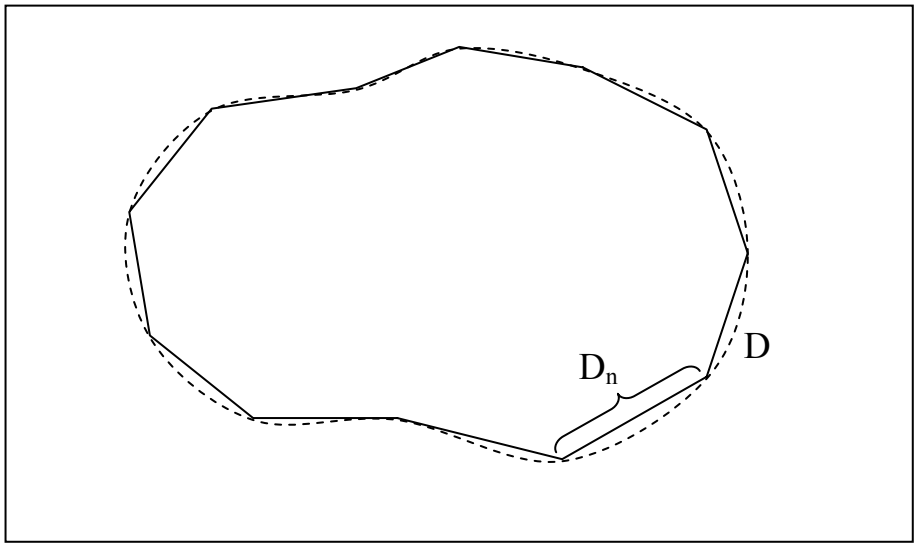


Figure 3.3 – Example of 2D domain segmentation.

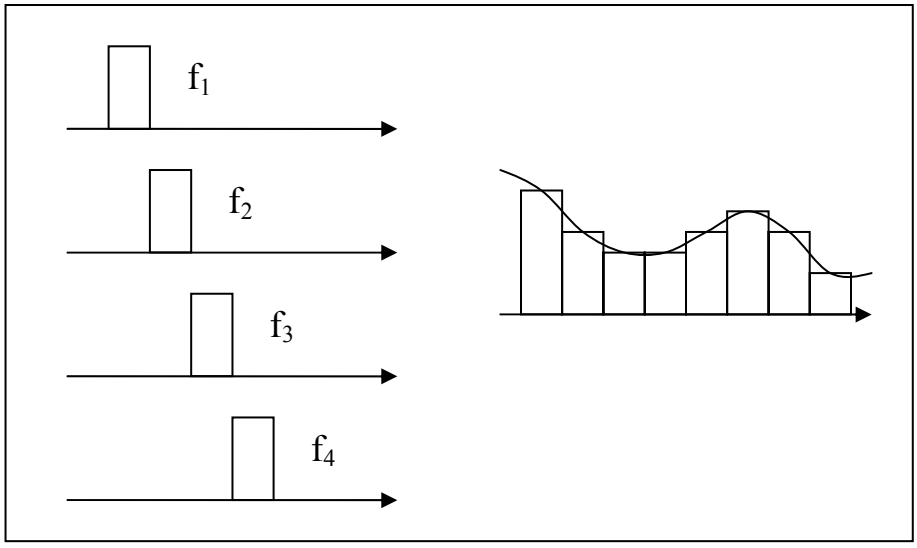


Figure 3.4 – Pulse basis functions and approximation of a continuous function.

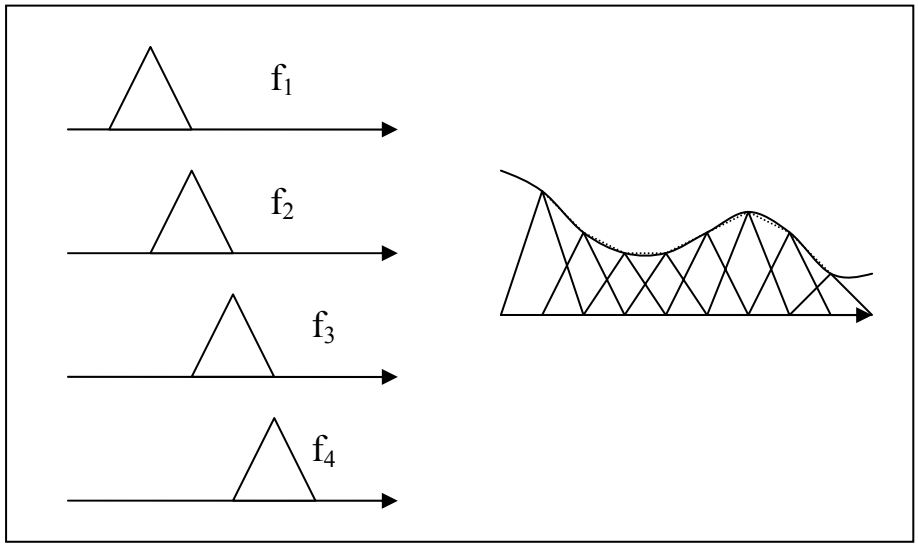


Figure 3.5 – Rooftop basis functions and approximation of a continuous function.

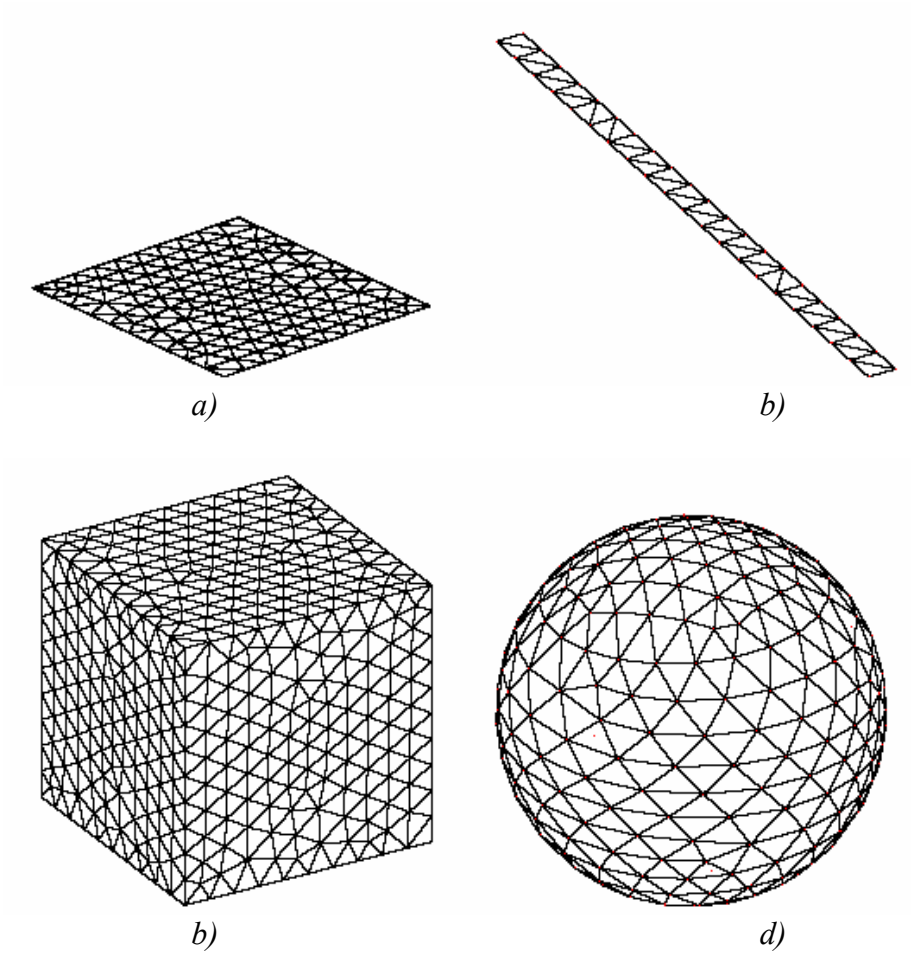


Figure 3.6 - Decomposition of a surface in triangular patches; a) square sheet; b) planar strip; c) cube; d) sphere.

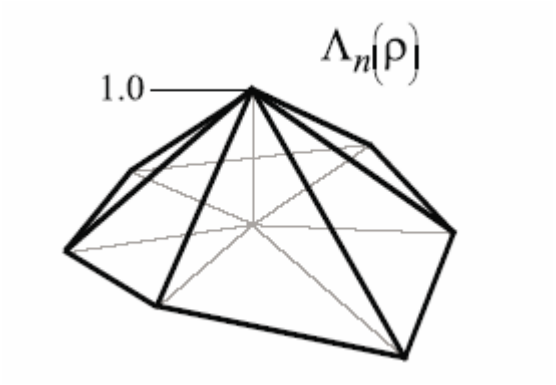


Figure 3.7 – Example of pyramidal basis function.

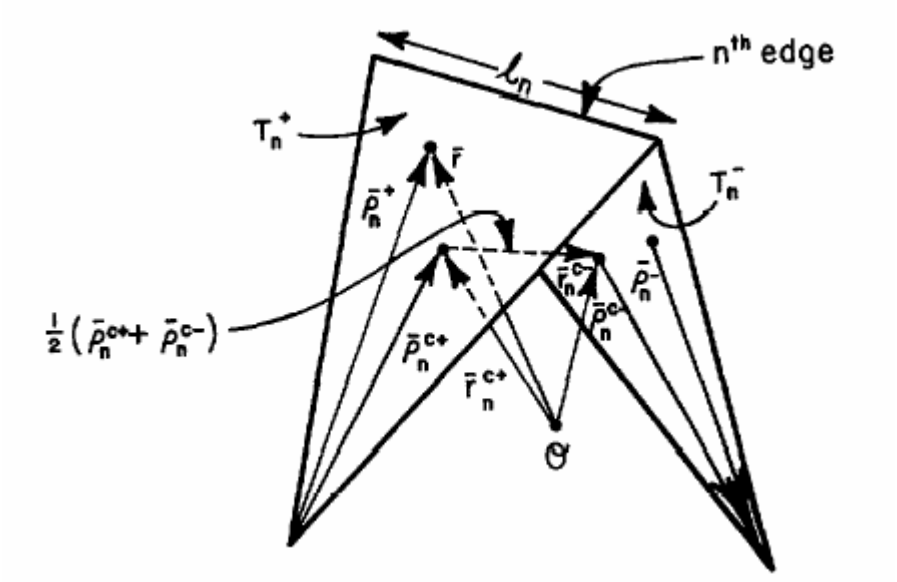


Figure 3.8 – Example of RWG basis function.

3.2 The MAS formulation

The scattering problem, apart from the classical Method of Moments, can be easily solved with the Method of Auxiliary Sources. While the MoM strategy aims to determine a distribution of equivalent currents on the boundary surface, the MAS formulation involves the use of discrete Auxiliary Sources (ASs) placed on a fictitious surface satisfying the correct boundary conditions on S. A popular choice is to use a set of elementary sources (filamentary currents in two-dimensional problems and elementary dipoles in three-dimensional ones) on a so-called auxiliary surface, placed just beneath the boundary surface: the fields radiated by such sources constitutes a non-orthogonal basis for the expression of the scattered field in the outer region.

3.2.1 MAS formulation in 2D

As done before for the MoM approach, let us consider a PEC object, its surface called *S*, immersed in free space and illuminated by and incident field E_i with a TM polarization (see fig. 3.9a).
Due to the 2D geometry and the TM polarization, all the electric fields are oriented along i_z, so that a scalar problem is considered.
Let us indicate with E_s the field scattered by the object in the region outside *S*. A set of elementary sources are placed on an auxiliary surface *S'*, conformal to *S,* located inside the object's volume (see fig. 3.9b).

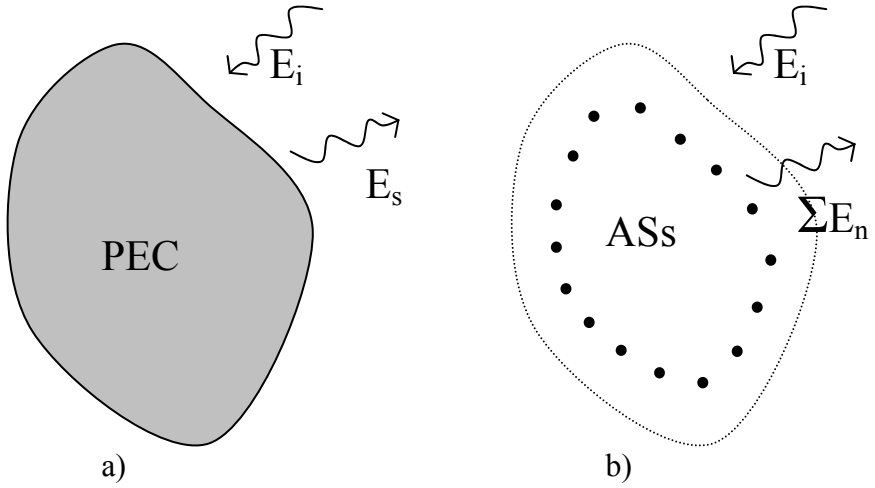


Figure 3.9 – a) problem geometry; b) MAS equivalent problem.

In order to evaluate the scattered field, the excitation of each source must be determined so that the superimposition of all the fields radiated in free space equals the scattered field. B
I.e., being E_n the field radiated by the n_th source, in the region outside S we must have:

$$\sum_{n=1}^{N} E_n = E_s \qquad (3.22)$$

Similarly to the MoM approach, in order to determine the source excitation, we have to enforce the boundary condition, that in the case of a PEC scattering object is written:

$$E_s|_S + E_i|_S = 0 \qquad (3.23)$$

The (3.23) gives us the value of the scattered field on the boundary S. The equivalence principle ensures that if the (3.22) is enforced on the

surface S, the field radiated by the set of ASs corresponds to the effective scattered field in the whole region outside S.
In any practical instance, the condition (3.23) is enforced on a number of so called collocation points (CPs) equal to the number of sources *N*. The collocation points are usually homogeneously distributed on the boundary *S*. In analogy with MoM, these play the role of impulsive testing functions, while the ASs plays the role of the basis functions.
The field radiated at the observation point $\underline{r}'$ by the n-th (filamentary) auxiliary source located at $\underline{r}_n$, can be written as:

$$E_n(\underline{r}')=a_n H_0^{(2)}(k|r_n-r'|) \tag{3.24}$$

where $H_0^{(2)}(\cdot)$ is the Hankel function of second kind and order zero, k is the propagation constant, and a_n is an excitation coefficient to be determined.
By enforcing the condition (3.23) on a set of *N* points on *S*, indicated as ($\underline{r}'_1 .. \underline{r}_m, \underline{r}'_N$), and using the field expression (3.24), a system of linear equations is found:

$$\underline{E}_i = \underline{\underline{G}}\,\underline{a} \tag{3.25}$$

where $\underline{E}_i$ is the vector whose components are the values of the incident field at the collocation points ($r'_1 .. r_m, r'_N$), and the *N*-by-*N* matrix $\underline{\underline{G}}$ expresses the interactions between the sources and the collocation points, i.e.,

$$G_{n,m}= H_0^{(2)}(k|r_n-r'_m|) \tag{3.26}$$

The vector $\underline{a}$, whose components are the excitation coefficients, is found by solving the system (3.25). Existence and uniqueness of the solution of the linear system arising from the MAS have been proved in [41]. With a proper choice of the CPs, the excitations obtained by solving the equation (3.25) ensure that the field radiated by the ASs is a good approximation of the actual scattered field on the whole domain *S*. Accordingly, on the basis of the equivalence principle, it can be assumed that also the fields radiated by the ASs is a good approximation of the scattered field in the volume outside *S*.

Details on the criteria underlying the choice of the CP and the AS are provided in the following.

3.2.2 Positioning the Auxiliary Sources and the testing points

The positioning of the Auxiliary Sources and testing points is a very critical issue of the MAS. As a matter of fact, general rules addressing the optimal positioning of the auxiliary sources and the collocation points are not available in the literature. The only available prescription requires that the auxiliary sources be located on a surface enclosing the singularities of the analytical continuation of the scattered field inside the scatterer body [42]. Usually, for smooth scatterers the auxiliary surface is chosen conformal to the boundary surface, and the ASs and the CPs are uniformly distributed. In the case of non smooth scatterers, two approaches are proposed. In the first one, the rate of AS and CP per unit length increases near the wedges, and the auxiliary surface is suitably shaped to approach the wedge vertexes. In the second approach, wedges are rounded out and eliminated from the scatterer profile, thus obtaining a smooth surface that approximates the original one [45, 46] (fig. 3.10).

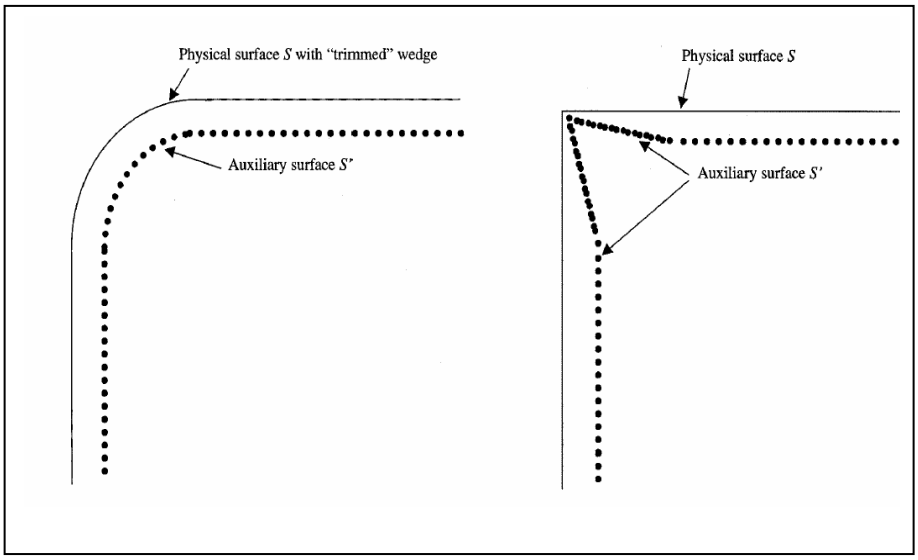


Figure 3.10 – Strategies for dealing with wedges: a) wedge rounding; b) thickening of the Auxiliary Sources.

Concerning the positioning of the collocation points on the boundary surface, we have to point out that the AS technique is essentially based on enforcing a relationship between two continuous functions (the incident and the scattered field on the boundary, see eq. (3.24)) on a finite number of points and on the assumption that this will ensure that the relationship holds (with a satisfactory accuracy) on the whole boundary surface. Obviously, the correctness of this assumption strictly depends on the spatial spectrum of the functions of interest as well as on the choice of the CPs.
The 'dept' of the auxiliary surface, i.e. the minimum distance between the ASs and the CPs, is an important parameter, since it strongly influences the performances of the technique. As pointed out by Anastassiu, Lymperopoulos and Kaklamani in a recent work [47], with reference to a 2D scatterer of circular shape, the evaluation error of the scattered field due to the MAS implementation decays when the 'dept' increase, but at the same time the displacement of the AS on an inner circumference of smaller radius causes the condition number to increase with an exponential law. Accordingly, when implementing a

MAS technique, a particular care should be used in choosing the dimension of the auxiliary surface and the number of AS, which determine the condition number of the impedance matrix, in order to be sure that the ill-conditioning will not affect the accuracy of the solution.

3.2.3 MAS formulation in 3D

In three-dimensional problems, the auxiliary sources are pairs of perpendicular elementary dipoles, both tangential to the surface *S'*. Analogously, pairs of point-wise testing functions are used to impose the boundary condition on the tangential components of the scattered and incident fields [48].
For example, with reference to a spherical scatterer of radius ρ_S, a grid is defined by choosing a set of *N* spherical directions (θ_n,φ_n), *n* from 1 to *N*.
A spherical auxiliary surface of radius ρ_A is defined, and in each point $r_{An}= (\rho_A ,\theta_n,\varphi_n)$ a pair of auxiliary sources are placed along the two versors $\hat{\theta}$, $\hat{\varphi}$ thus. the impulsive current distributions:

$$\underline{J}'_n(r) = \delta(r_n)\,\hat{\theta}\,; \quad \underline{J}''_n(r) = \delta(r_n)\,\hat{\varphi} \tag{3.22}$$

The boundary condition are enforced by considering the tangential components of the incident and the ASs' field on a set of points $r_n=(\rho_S ,\theta_n,\varphi_n)$ n=1,…,N, on the scattering surface.
We have two sets of testing functions:

$$\underline{w}'_n(r) = \delta(r_n)\,\hat{\theta}\,; \quad \underline{w}''_n(r) = \delta(r_n)\,\hat{\varphi} \tag{3.23}$$

For the sake of convenience often the ASs $\underline{J}'_{1..N}$ and $\underline{J}''_{1..N}$ are merged together considering a single set of ASs currents, say $\underline{J}_{1..2N}$. Analogously, the testing functions usually are merged together considering a single set, say $\underline{w}_{1..2N}$.

The field radiated by each AS is evaluated by means of the 3D Green's function in free space. Analogously to the 2D case, a linear system relating the unknown values $\underline{J}_{1..2N}$ to the known incident field is found and solved.

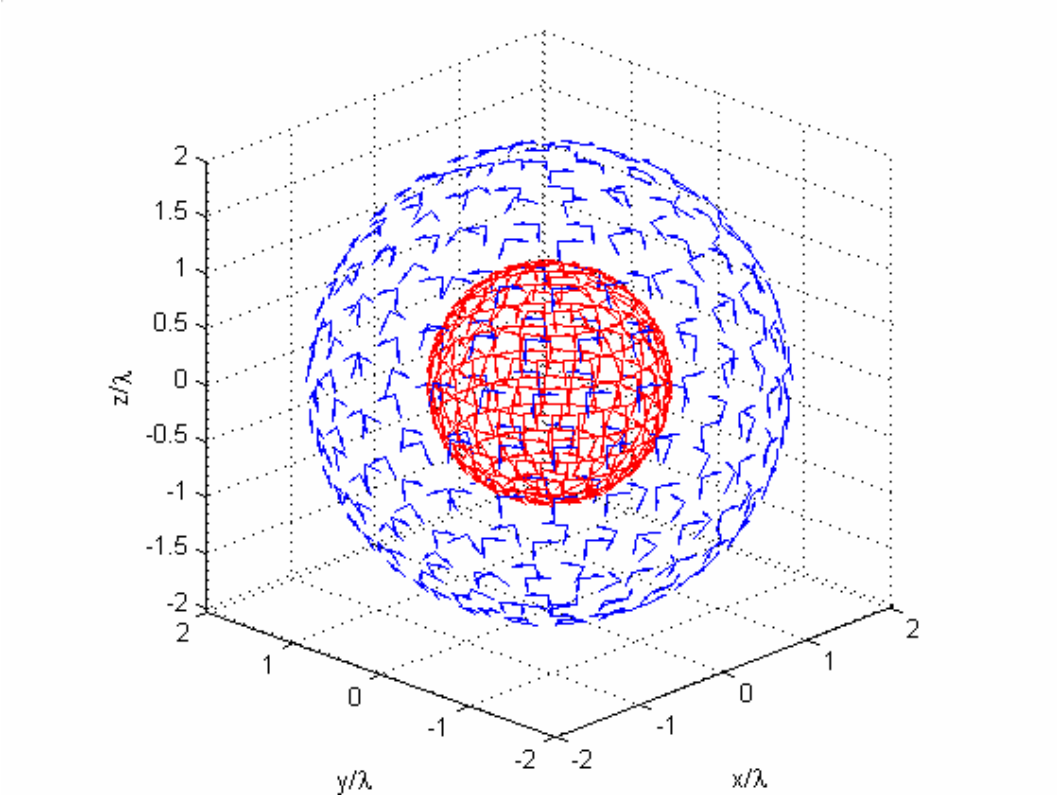


Figure 3.10 – Positioning of auxiliary sources and testing functions

It is interesting to note that the MAS can be of particularly interest in the 3D case, where the classical MoM techniques are scarcely used due to the excessive number of unknowns.

3.3 MAS advantages and disadvantages with respect to the standard MoM

The application of MAS to the analysis of large metallic scatterers enjoy significant advantages with respect to the classical MoM-based techniques.
The first one is related to the kind of basis functions adopted. In fact, in, the set of basis functions the MoM-based techniques should be able to represent any possible continuous current distribution induced by the incident field. For large scatterers, this leads to basis functions (rooftop, RWG, …), continuous on a sub-domain of the boundary surface. In order to numerically evaluate the radiated field, the current distribution of each function has to be discredited by means certain number of points, and all the contributions are then summated. Concerning the testing, the Galerkin method is often adopted, that consist in choosing the testing functions coincident with the current expansion functions. For the same purpose, simple pointwise functions are instead adopted in the MAS ("collocation method"), thus reducing the computational effort in enforcing the boundary condition. Another problem pertains to the evaluation of the "self-terms" of the impedance matrix which is a critical issue in MoM techniques and has to be properly addressed. In fact, their evaluation involve the singularity of the Green function for argument equal to zero.
In the MAS approach, it can be said that the role of the basis functions is played by a set of elementary point-wise sources located inside the scatterer surface. The field radiated on the boundary (and elsewhere) by each source is analytically known and can be very easily evaluated. At variance of the M.o.M. no discretizations and integrations are needed. Furthermore, the self-terms evaluation problem is avoided, since the sources are located inside the scatterer domain and the Green function is always evaluated for arguments away from the singularity. However, the main advantages of the MAS consists in the capability of studying the scattering by a large object by means of a very small number of unknowns. In fact, it will be shown that an optimized MAS technique can solve this problems by means of a number of AS and testing function very close to the number of degrees of freedom of the scattered field (i.e. 2 per wavelength), while a standard MoM

technique, to attain the same accuracy requires at least 10 unknowns per wavelength [20].

On the other hand, to give a comprehensive understanding of the MAS capabilities, some critical points concerning its implementation are to be illustrated, and some possible solutions proposed.

The first negative issue concerns the increase of the condition number of the impedance matrix, due to the increase of the distance of the AS from the scatterer surface. However, it must be noted that this problem become prohibitive only when the precision of the available numerical tools is not high enough to manage the ill conditioning of the problem.

As previously noted, the most critical issue of the MAS approach is related to the positioning of the Auxiliary Sources and Collocation Points. This problem will be addressed in this thesis by exploiting the theory of the spatial bandwidth of the scattered fields introduced by Bucci and Franceschetti [2, 50]. In particular, starting from the optimal sampling theory [51], we will formulate some useful guidelines for the optimal implementation of the MAS, strongly improving its performances, and proposing it as a non-redundant technique for the analysis of large scatterers.

Chapter 4

The Fast Multipole Method

An innovative approach to the boundary value problem, called "Fast Multipole Method", was proposed by Greengard and Rokhlin for the evaluation of the potential and force fields in a system involving a large number of particles [36], starting from Rokhlin's work on the fast solution of the Laplace equation [37]. Subsequently, the FMM approach has been successfully extended to the Helmholtz equation arising from the acoustic scattering problem [38].

In recent years, the FMM has been applied to many other issues involving the multiplication of N by N matrices by vector of length N, and it gained a place among the best ten algorithms of the 20th Century [52].

A field of particular interest for the application of the FMM paradigm the Computational Electromagnetics. As discussed in the previous chapters, even if the available computational capacity undergoes a continuous increase, the interest in the full-wave analysis of large scattering and radiating structures still involves problems of prohibitive size from the numerical point of view.

The solution of the system of linear equations arising from an integral formulation (for example, a MoM technique) of the scattering problem is usually obtained by means of an iterative approach, in order to reduce the computational complexity of the matrix inversion (from $O(N^3)$, required by a direct solver) to $O(N^2)$. The FMM paradigm can be applied to the matrix-by-vector multiplications entailed in the iterative solver, reducing the overall computational complexity to $O(N^{3/2})$ or lower. In particular, the recursive application of the FMM over a multilevel scheme, that goes under the name of MultiLevel Fast Multipole Method (MLFMM) [39, 40], is capable to reduce the numerical effort up to $O(N\log N)$.

4.1 Iterative Methods

The solution of many scientific problems require the solution of large linear systems in the form:

$$\underline{\underline{A}}\,\underline{x} = \underline{b} \tag{4.1}$$

where $\underline{\underline{A}}$ is an N-by-N matrix, $\underline{b}$ is a known vector of length N, and x is the solution to determine. The direct inversion of the matrix $\underline{\underline{A}}$ involves the application of procedures like the Gaussian elimination or the LU factorization, that requires a number of operations proportional to N^3. In most application such computational complexity is prohibitive, so that the problem cannot be solved in a straight way. As an alternative, many iterative approaches have been proposed to solve approximately the linear system (1). They allows to evaluate a sequence $\{\underline{x}_n\}$ converging to the solution of (1) so that a good approximation is reached after a suitable finite number of steps. Each iteration essentially requires a matrix-by-vector product and involves a computational complexity of $O(N^2)$. Accordingly, the overall numerical effort is reduced, provided that a satisfactory estimate of the solution is obtained within a number of steps independent on N.

Iterative methods can be subdivided in two classes: stationary and nonstationary methods. Stationary methods, like the Jacobi method, the Gauss Seidel method, the Successive Overrelaxion method (SOR) [6], were the first to be proposed and applied, while the nonstationary methods, like the Conjugate Gradient (CG), the Biconjugate Gradient (BiCG) and the Generalized Minimal Residual (GMRES)[53], are relatively recent.

Nonstationary methods are based on the idea of sequences of orthogonal vectors, and differ from the stationary ones because the information involved by the computation changes from one iteration to the other. Nonstationary methods are more complicated and harder to be implemented, but they are typically more effective than the stationary ones [6]. Even if there is not an universally agreed definition, they're often denoted as Krylov Methods [54].

A Krylov method starts with an initial guess x_0 and performs a series of improving iterations. At k_th iteration, the method produces an approximate solution $\underline{x}_k$ from a Krylov space generated by a vector $\underline{c}$:

$$K_k\left(\underline{\underline{A}},\underline{c}\right)=span\left\{\underline{c},\underline{\underline{A}}\underline{c},...,\underline{\underline{A}}^{k-1}\underline{c}\right\} \qquad (4.2)$$

where $\underline{\underline{A}}^n\underline{c}$ is the result of n applications of the operator $\underline{\underline{A}}$ to the vector $\underline{c}$. A common choice for the initial condition is $\underline{x}_0=\underline{0}$ and $\underline{c}=\underline{b}$.
In the following, some popular Krylov methods are briefly illustrated.

4.1.1 The Conjugate Gradient Method

Developed by Hestenes and Stiefel [55], the Conjugate Gradient method has been the first proposed nonstationary method and has been intensively used because of its effectiveness.
The CG method produces a sequence of iterates $\underline{x}_1,...,\underline{x}_N$, such that the norm of the difference between the $\underline{x}(k)$ at the k-th step and the exact solution, is minimized. At k-th step, the iterate $\underline{x}(k)$ belongs to the space

$$K_k\left(\underline{\underline{A}},\underline{r}^{(0)}\right)=span\left\{\underline{r}^{(0)},\underline{\underline{A}}\underline{r}^{(0)},...,\underline{\underline{A}}^{k-1}\underline{r}^{(0)}\right\} \qquad (4.3)$$

where $\underline{r}^{(0)}$ is the residual relative to the initial guess $\underline{x}^{(0)}$: $\underline{r}^{(0)}=\underline{b}-\underline{\underline{A}}\underline{x}^{(0)}$.
The method proceeds as follows:

$for\ k=1,2,3...$

$$\rho_{k-1}=\underline{r}^{(k-1)^T}\underline{r}^{(k-1)}$$

$$\beta_{k-1}=\begin{cases}0 & if\ k=1\\ \rho_{k-1}/\rho_{k-2} & if\ k\neq 1\end{cases}$$

$$\underline{p}^{(k)}=\underline{r}^{(k-1)}+\beta_{k-1}\underline{p}^{(k-1)}$$

$$\underline{q}^{(k)}=\underline{\underline{A}}\underline{p}^{(k)}$$

$$\alpha_k = \rho_{k-1} \Big/ \underline{p}^{(k)^T} \underline{q}^{(k)}$$
$$\underline{x}^{(k)} = \underline{x}^{(k-1)} + \alpha_k \underline{p}^{(k)}$$
$$\underline{r}^{(k)} = \underline{r}^{(k-1)} - \alpha_k \underline{q}^{(k)}$$

end

where the apex *T* indicates the transpose of the column vectors.
At each iteration, the CG generates a tentative solution vector, a residual vector, and a search direction used to update them.
Starting from the bottom lines, it can be seen that the solution $\underline{x}^{(k)}$ is updated at every step with a correction in the direction $\underline{p}^{(k)}$, while the residual is modified by a correction in the direction $\underline{q}^{(k)}$. It is possible to show that the choice of the scalars α_k and β_k makes the search direction $\underline{p}^{(k)}$ and the residual $\underline{r}^{(k)}$ orthogonal to all previous $\underline{\underline{A}}\underline{p}^{(i<k)}$ and $\underline{r}^{(i<k)}$, respectively.
According to these considerations, the CG algorithm requires at each step only one matrix-by-vector product, two inner product and three vector updates. From a computational point of view, the overall complexity is dominated by the matrix-by-vector multiplication, of the order $O(N^2)$, N being the size of the square matrix. Moreover, the method just requires the storage of a few number of vectors.
Some guidelines can be given on the convergence properties of the method, even if an accurate prediction is not possible. The behaviour of the approximation error as a function of the number of iterations is strongly related to the condition number of the matrix A, say $\kappa(A)$. In fact, it can be shown that:

$$\left\| \underline{x}(i) - \underline{\hat{x}} \right\|_A \leq 2\varepsilon^i \left\| \underline{x}^{(0)} - \underline{\hat{x}} \right\|_A \qquad (4.4)$$

where $\underline{\hat{x}}$ is the exact solution, $\varepsilon = \left(\sqrt{\kappa} - 1\right) \Big/ \left(\sqrt{\kappa} + 1\right)$, and $\left\| \underline{y} \right\|_A = (\underline{y}, \underline{\underline{A}}\underline{y})$. This equation shows that in order to have a bounded error it is necessary to perform a number of iteration proportional to $\sqrt{\kappa}$.
The use of a proper preconditioning matrix *M*, usually tailored on the specific problem under consideration, helps to ameliorate the

convergence properties. In particular, the preconditioner should be such that the matrix $\underline{\underline{M}}\cdot\underline{\underline{A}}$ should have a lower condition number with respect to the original matrix $\underline{\underline{A}}$.

4.1.2 The Bi-Conjugate Gradient Method

Proposed by Fletcher [56], the Bi-Conjugate Gradient Method (BiCG) is an improvement of the classical CG scheme, useful when the matrix is nonsymmetric and non-singular [57]. The key idea behind the BiCG is to replace the orthogonal sequence of residuals with two mutually orthogonal sequences, by using both the original matrix $\underline{\underline{A}}$ and its transposed $\underline{\underline{A}}^T$. Analogously to the CG, an initial guess $\underline{x}^{(0)}$ is chosen (often equal to zero), but two sequences of residuals are needed, say $\underline{r}^{(k)}$ and $\underline{\tilde{r}}^{(k)}$, instead of one. The starting value of the first residual is chosen like in CG, i.e. $\underline{r}^{(0)} = \underline{b} - \underline{\underline{A}}\underline{x}^{(0)}$, while the second choice is free, even if usually is set $\underline{\tilde{r}}^{(0)} = \underline{r}^{(0)}$. The BiCG finds an approximate solution to the linear system in the space:

$$K_k\left(\underline{\underline{A}}, \underline{r}^{(0)}\right) = span\left\{\underline{r}^{(0)}, \underline{\underline{A}}\underline{r}^{(0)}, ..., \underline{\underline{A}}^{k-1}\underline{r}^{(0)}\right\} \qquad (4.5)$$

and orthogonally to the space:

$$L_k\left(\underline{\underline{A}}, \underline{\tilde{r}}^{(0)}\right) = span\left\{\underline{\tilde{r}}^{(0)}, \underline{\underline{A}}^T\underline{\tilde{r}}^{(0)}, ..., \underline{\underline{A}}^{Tk-1}\underline{\tilde{r}}^{(0)}\right\} \qquad (4.6)$$

The method proceeds according to the following scheme:

$for\ k = 1,2,3...$

$$\rho_{k-1} = \underline{r}^{(k-1)^T}\underline{\tilde{r}}^{(k-1)}$$

$$\beta_{k-1} = \begin{cases} 0 & if\ k = 1 \\ \rho_{k-1}/\rho_{k-2} & if\ k \neq 1 \end{cases}$$

$$\underline{p}^{(k)} = \underline{r}^{(k-1)} + \beta_{k-1}\underline{p}^{(k-1)}$$

$$\underline{\tilde{p}}^{(k)} = \underline{\tilde{r}}^{(k-1)} + \beta_{k-1}\,\underline{\tilde{p}}^{(k-1)}$$

$$\underline{q}^{(k)} = \underline{\underline{A}}\,\underline{p}^{(k)}$$
$$\underline{\tilde{q}}^{(k)} = \underline{\underline{A}}^{T}\,\underline{\tilde{p}}^{(k)}$$

$$\alpha_k = \rho_{k-1} \Big/ \underline{\tilde{p}}^{(k)^T}\,\underline{q}^{(k)}$$
$$\underline{x}^{(k)} = \underline{x}^{(k-1)} + \alpha_k\,\underline{p}^{(k)}$$
$$\underline{r}^{(k)} = \underline{r}^{(k-1)} - \alpha_k\,\underline{q}^{(k)}$$
$$\underline{\tilde{r}}^{(k)} = \underline{\tilde{r}}^{(k-1)} - \alpha_k\,\underline{\tilde{q}}^{(k)}$$

end

Together with the residual sequence $\underline{\tilde{r}}^{(k)}$, also a search direction vector sequence is added, $\underline{\tilde{p}}^{(k)}$. The choice of the scalars α_k and β_k ensures the bi-orthogonality relations:

$$\begin{aligned} \underline{\tilde{r}}^{(k)^T}\,\underline{r}^{(h)} &= 0 \\ \underline{\tilde{p}}^{(k)^T}\,\underline{\underline{A}}\,\underline{p}^{(h)} &= 0 \end{aligned} \qquad \textit{for } k \neq h \qquad (4.7)$$

The BiCG has been applied with good results to non-symmetrical systems. Concerning the convergence properties, sometimes an irregular behaviour has been denoted: typical break down situation can occur when $\underline{r}^{(k-1)^T}\,\underline{\tilde{r}}^{(k-1)} \approx 0$, or when $\underline{\tilde{p}}^{(k)^T}\,\underline{q}^{(k)} \approx 0$.

4.1.3 The Generalized Minimal Residual method

Proposed by Saad and Schultz in 1986 [53], the Generalized Minimal Residual method (GMRES) is useful to solve general nonsymmetric systems.

The GMRES method founds the best solution $\underline{x}_k$ from the Krylov space $K_k\left(\underline{\underline{A}},\underline{r}^{(0)}\right)$. This means that, at each step, $\underline{x}_k$ solves the least square problem:

$$\min_{\underline{x}\in K_k\left(\underline{\underline{A}},\underline{r}^{(0)}\right)}\left\|\underline{b}-\underline{\underline{A}}\underline{x}\right\| \tag{4.8}$$

where ||.|| is the Euclidean norm.
In the GMRES, an orthonormal basis $\left\{\underline{v}_1,...,\underline{v}_k\right\}$ for the space $K_k\left(\underline{\underline{A}},\underline{r}^{(0)}\right)$ is constructed explicitly by means of the Arnoldi's method [58], which is a version of the Graham-Schmidt procedure. Such basis is retained and updated at every step during the iterative process.
The method proceeds according to the following scheme:

$$\underline{x}_0 = 0, \quad \underline{v}_1 = \underline{b}\Big/\left\|\underline{b}\right\|, \quad \underline{\underline{V}}_1 = \underline{v}_1$$

for $k = 1,2,3...$

$$\underline{h}_k = \underline{\underline{V}}_k^{*}\,\underline{\underline{A}}\underline{v}_k$$

$$\hat{\underline{v}}_{k+1} = \underline{\underline{A}}\underline{v}_k - \underline{\underline{V}}_k\,\underline{h}_k$$

$$\underline{v}_{k+1} = \left(\hat{\underline{v}}_{k+1}\right)\Big/\left\|\hat{\underline{v}}_{k+1}\right\|$$

$$\underline{\underline{V}}_{k+1} = \left[\underline{\underline{V}}_k \quad \underline{v}_{k+1}\right]$$

$$\underline{\underline{H}}_k = \begin{bmatrix} \underline{\underline{H}}_{k-1} & \underline{h}_k \\ 0 & \left\|\hat{\underline{v}}_{k+1}\right\| \end{bmatrix}$$

$$\underline{y}_k = \min_{\underline{y}}\left\|\begin{bmatrix} \left\|\underline{b}\right\| \\ 0 \\ ... \\ 0 \end{bmatrix} - \underline{\underline{H}}_k\,\underline{y}\right\|$$

$$\underline{x}_k = \underline{\underline{V}}_k\,\underline{y}_k$$

end

The first vector of the basis is obtained by normalizing the known term vector b, which constitutes the base for the space $K_1(\underline{\underline{A}},\underline{b})$. Then, a procedure evaluating an orthonormal basis for $K_{k+1}(\underline{\underline{A}},\underline{b})$ from $K_k(\underline{\underline{A}},\underline{b})$ is applied recursively. At each step, the $(k+1)$_th basis vector $\underline{v}_{k+1}$ is obtained by othornormalizing the vector $\underline{\underline{A}}\underline{v}_k$ against the subspace $K_k(\underline{\underline{A}},\underline{b})$. The basis vectors are stored and retained in the matrix V_k. The least square problem (8) is solved by projecting the solution on the orthogonal basis, providing at each step the solution $\underline{x}_k = \underline{V}_k\underline{y}_k$.
In principle, the method provides the exact solution after N iterations. However the procedure is usually terminated as soon as the residual becomes small enough.
To reduce the storage requirements, the method is often implemented in "restarted" version: after m iteration, all the retained data are cleared up, and the procedure is restarted from the last found tentative solution.

4.2 Preconditioning

As previously mentioned, the convergence rate of the iterative methods depends on the spectral properties of the coefficient matrix. When the matrix representing the linear system to solve is scarcely conditioned, the number of iteration required to reach an acceptable approximation of the solution can be prohibitive. As pointed out by the relation (4.4), the precision achievable after a finite number of steps is related to the condition number of the matrix.
The countermeasure set up by the researchers consists in turning the original linear system into a better-conditioned one, by means of a so-called preconditioning technique [5]. A preconditioner is a matrix that transforms the original linear system (4.1) into:

$$\underline{\underline{M}}^{-1}\underline{\underline{A}}\underline{x} = \underline{\underline{M}}^{-1}\underline{b} \qquad \text{(left preconditioning)} \qquad (4.9)$$

or:

$$\underline{\underline{A}}\underline{\underline{M}}^{-1}(\underline{\underline{M}}\underline{x}) = \underline{b} \qquad \text{(right preconditioning)} \qquad (4.10)$$

Obviously, the systems (4.9) and (4.10) have the same solution than the original system, although the matrices $\underline{\underline{M}}^{-1}\underline{\underline{A}}$ and $\underline{\underline{A}}\underline{\underline{M}}^{-1}$, in principle, can show spectral properties different than the original one. The aim of a preconditioning system is to find a matrix $\underline{\underline{M}}$ ensuring a faster convergence of the iterative procedures due to the improved spectral properties of the matrix $\underline{\underline{M}}^{-1}\underline{\underline{A}}$ or $\underline{\underline{A}}\underline{\underline{M}}^{-1}$.

Generally speaking, all the iterative methods discussed above suffer from slow convergence and can be improved by a suitable preconditioning.

In the case of techniques useful only to symmetric systems, like the CG one, a little complication can arise. In fact, even if M is a symmetrical matrix, $\underline{\underline{M}}^{-1}\underline{\underline{A}}$ or $\underline{\underline{A}}\underline{\underline{M}}^{-1}$ can be nonsymmetric, and the solution of the transformed system is not possible. A simple method to solve this problem is based on the splitting of the preconditioner as a product of two matrices, i.e.,:

$$M = \underline{\underline{M}}_1 \underline{\underline{M}}_2 \tag{4.11}$$

where $\underline{\underline{M}}_2 = \underline{\underline{M}}_1^{\mathrm{T}}$, and then on the simultaneous application of the left and right conditioning, i.e.,:

$$\underline{\underline{M}}_1^{-1}\underline{\underline{A}}\underline{\underline{M}}_2^{-1}(\underline{\underline{M}}_2\underline{x}) = \underline{\underline{M}}_1^{-1}\underline{b} \tag{4.12}$$

In this way, the matrix $\underline{\underline{M}}_1^{-1}\underline{\underline{A}}\underline{\underline{M}}_2^{-1}$ is symmetrical an the CG method can be applied.

Most of the preconditioners try to find a matrix M that, in some sense, can approximate the matrix A so that, with reference to the formulation (4.9), $\underline{\underline{M}}^{-1}\underline{\underline{A}}$ is close to an identity matrix, and can be quite easily inverted.

Preconditioner can be derived from the knowledge of the original physical problem from which the linear system arise, or can be constructed directly from the data contained in the original matrix.

In the following, some popular approaches are described.

4.2.1 Diagonal Preconditioning

The simplest preconditioning technique consists in taking the matrix $\underline{\underline{M}}$ equal to the diagonal of the original impedance matrix $\underline{\underline{A}}$:

$$m_{i,j} = \begin{cases} a_{i,j} & for\ i = j \\ 0 & for\ i \neq j \end{cases}$$

where $a_{i,j}$ is the generic element of the matrix $\underline{\underline{A}}$. This approach, even if rough, is quite inexpensive from a computational point of view, and can be effective in reducing the condition number when the original matrix is quite diagonal.

4.2.2 Block Jacobi Preconditioning

The block Jacobi preconditioning is based on a partitioning of the variable space. The set of the indexes $S = \{1,...,N\}$ is partitioned into a collection of H disjoint subsets $S_{1..H}$: $S = \bigcup_{h=1}^{H} S_h$. The preconditioner $\underline{\underline{M}}$ is built as:

$$m_{i,j} = \begin{cases} a_{i,j} & if\ i\ and\ j\ belong\ to\ the\ same\ subset \\ 0 & otherwise \end{cases}$$

Obviously, that the diagonal preconditioning can be considered a particular case of block preconditioner.
Usually, the index partitioning arise from considerations on the physical problem at hand. For example, when dealing with the MoM and other IF-based methods for the electromagnetic scattering evaluation, the blocks arise from the partitioning of the boundary surface and the consequent clustering of the expansion and testing functions defined on each partition.
From a computational point of view, this techniques implies the direct inversion of the block diagonal matrix, in order to evaluate $\underline{\underline{M}}^{-1}$. Such task correspond to the inversion of H square matrices of

dimension N/H, requiring an $O(N^3/H^2)$ complexity. Provided that the number of blocks is proportional to the root square of the number of unknowns, a popular choice, the computational effort becomes $O(N^{1.5})$. Thus, the application of this preconditioning technique can be considered fairly inexpensive.
The block Jacobi preconditioner is effective in those cases where the matrix can be rearranged so that the dominant terms are clustered around the diagonal.

4.2.3 Preconditioning based on the incomplete LU Factorization

As said previously, a direct inversion of the matrix is possible by means of a LU factorization, that requires a number of operation proportional to N^3, in most cases a prohibitive complexity. A class of techniques exploits the decomposition concept to produce an approximated version of the LU factorization by neglecting some elements of the matrix to invert [6]. In this way, a preconditioner $\underline{\underline{M}} = \underline{\underline{L}}\,\underline{\underline{U}}$ is obtained, such that $\underline{\underline{M}}^{-1}$ approximates $\underline{\underline{A}}^{-1}$. In computational terms, the number of quasi-zero elements to neglect must be large enough to ensure significant savings with respect to the direct inversion.

4.2.4 Preconditioning based on the near interaction data in FMM

In the implementation of the Fast Multipole Method, some complicacies for the choice of the preconditioning arise from the fact that the impedance matrix is never computed explicitly. Anyway, the near field interactions, that fill the full-rank blocks of the matrix, are evaluated explicitly and constitute a good starting set of data to build a preconditioner that approximates $\underline{\underline{A}}$. In fact, it is clear that both the diagonal and the block-diagonal preconditioners can be easily implemented [59]. Moreover, also the ILU preconditioning technique can be applied to the FMM [60].

Anyway, in the most common approach the sparse matrix based on the NF interactions of the FMM is directly used as the preconditioner, as it can be considered an approximation of the impedance matrix.

4.3 The FMM strategy

The Fast Multipole Method and its various variants can be effectively applied to the MoM-based analysis of the scattering by arbitrary geometries [33]. The bottleneck of common MoM implementation is the iterative algorithm (CG, BiCG, GMRES, …) used for the inversion of the impedance matrix. As seen, it requires the evaluation at each step of the field produced by a tentative solution representing a current distribution, which involves a matrix-by-vector product that requires a number of operations proportional to N^2, N being the number of basis functions in use.
The impedance matrix is dense, because the interaction between any current expansion and any testing function is never zero, even if they're localized far away from each other. From a physical point of view, the evaluation, storage and usage of the full impedance matrix implies that all the interactions between any current expansion function and any testing function are taken into account singularly (see fig. 4.1a). The aim of the FMM approach is to turn this "full connected" network into a properly built scheme where the number of total links is reduced by means of "hubs" (see fig. 4.1b).

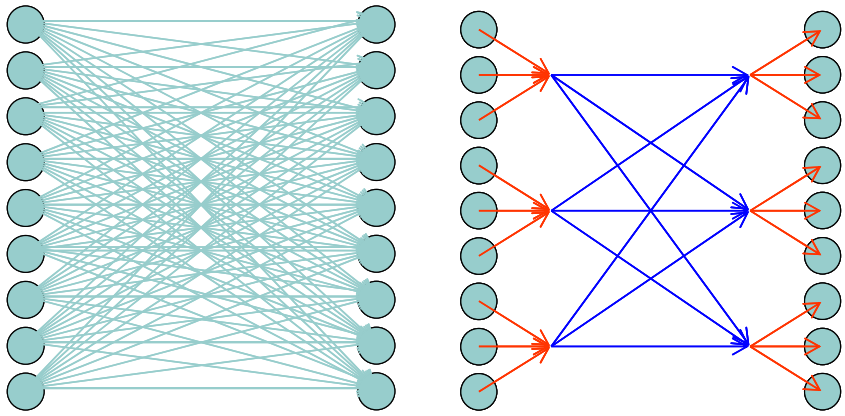


Figure 4.1 - a) Full connected network; b) Network exploiting "hubs" to reduce the number of links

The idea behind the FMM is that most of the elements of the impedance matrix, which represent the interaction between basis and testing function localized very far from each other, effectively involve very few degrees of freedom and carry redundant information. The strategy is to subdivide the basis and testing function into localized groups, and to "aggregate" the contributions due to the elements in each group, then to study the interactions between the couples of non-adjacent groups (fig. 4.2).
In this way, the N-by-N impedance matrix $\underline{\underline{Z}}$ is split into the sum of two matrices:

$$\underline{\underline{Z}} = \underline{\underline{Z}}' + \underline{\underline{Z}}'' \qquad (4.13)$$

$\underline{\underline{Z}}'$ taking into account the interaction between nearby basis and testing function, $\underline{\underline{Z}}''$ containing the blocks that expresses the interaction between non adjacent groups (see figure 4.3).

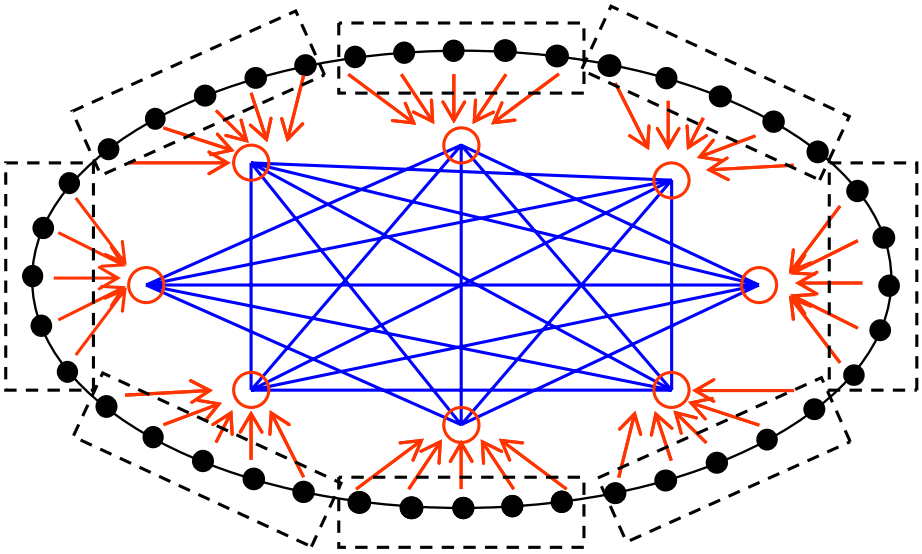


Figure 4.2 – Scheme representing the evaluation of the far interaction by means of the FMM. The black circles represents the nodes on the geometry, the red circles indicate the multipole centers. Relation between each node and its multipole center is indicated by a red arrow, interactions between the multipoles are depicted by blue lines.

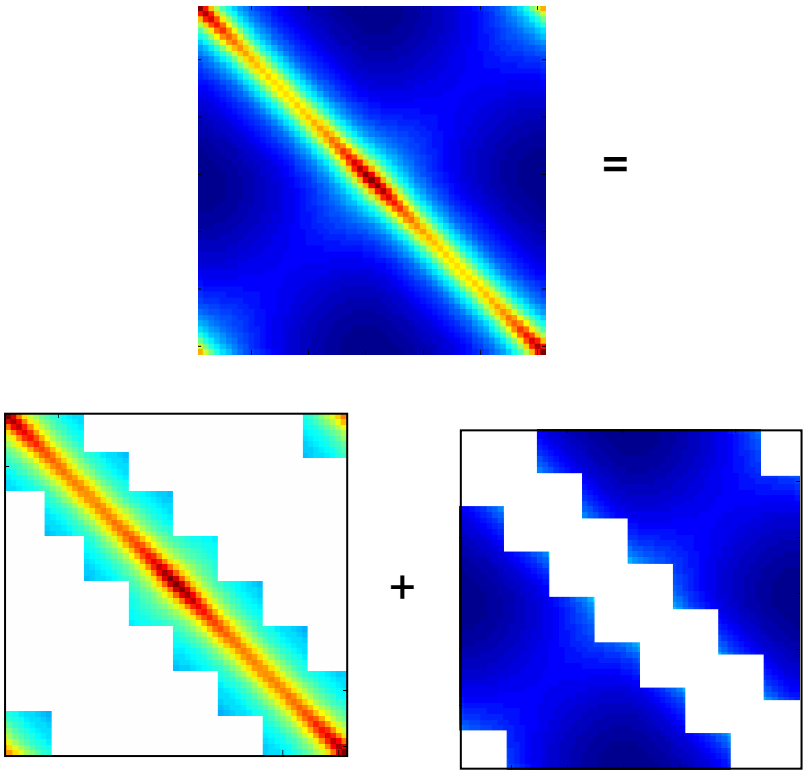


Figure 4.3 – The impedance matrix can be split in the sum of two matrices: $\underline{\underline{Z}} = \underline{\underline{Z}}' + \underline{\underline{Z}}''$.

From a mathematical point of view, it can be seen that the blocks (usually off-diagonal) constituting the matrix $\underline{\underline{Z}}''$ are practically rank deficient, thus can be conveniently expressed by means of a suitable low-redundancy representation.
To figure out how it is possible, it's interesting to note that an M-by-M (sub-) matrix $\underline{\underline{B}}$ of rank R can be expressed by means of a sum of outer products of vectors of length M:

$$\underline{\underline{B}} = \sum_{r=1}^{R} t_r \underline{u_r} \cdot \underline{v_r}^T \qquad (4.14)$$

This means that the *M*-by-*M* matrix $\underline{\underline{B}}$ can be stored by means of *R* vectors, with an occupancy proportional to *2RM* instead of M^2. Moreover, the matrix-by-vector product $\underline{\underline{B}}\underline{x}$, that usually requires N^2 operations, is carried out by means of a number of operations proportional to *RM*. In fact, the product

$$\underline{\underline{B}}\underline{x} = \sum_{r=1}^{R} \underline{u_r} \cdot t_r \cdot \underline{v_r}^T \cdot \underline{x} \tag{4.15}$$

for any fixed *r* involves three steps:

- the vector-by-vector product $\underline{v_r}^T \cdot \underline{x}$, called "aggregation", because in some sense it "joins" the contributes from the *M* elements of the vector $\underline{x}$ into a scalar result
- the multiplication of the above result by the scalar t_r, representing a "translation";
- the product of the translated value by the vector $\underline{u_r}$, named "disaggregation", since it derives the *M* elements of the resulting vector from the scalar result of the above step

It is clear that a number of operation proportional to the block size *M* is needed to carry out the three-step sequence, that is repeated *R* times.
When dealing with EM scattering, the required factorization can be obtained by applying the addition theorem to the Green's function; this leads to an analogous decomposition of the impedance matrix.
The scheme seen for the block *B* can be applied to the whole matrix containing the low rank block accounting for the interaction between non adjacent clusters, $\underline{\underline{Z}}$''. It is found that it can be written:

$$\underline{\underline{Z}}'' = \sum_{r=1}^{R} \underline{\underline{U}}_r \cdot \underline{\underline{T}}_r \cdot \underline{\underline{V}}_r^T \tag{4.16}$$

Where the matrices $\underline{\underline{V}}$ contain the aggregation vectors, $\underline{\underline{U}}$ the disaggregation vectors, and $\underline{\underline{T}}$ expresses the "translation" operator, i.e. the interaction between the clusters.
The group dimension *M* is chosen equal to $\sqrt{N}$, and it is also found *R* to be proportional to $\sqrt{N}$. We have to store *R* matrices $T_{1..R}$, of size

M-by-*M*, with an occupancy proportional to $N^{1.5}$, while the sparse matrices $V_{1..R}$ contain $N^{1.5}$ non-zero elements, and the same for the sparse matrices $U_{1..R}$. Again, it is found that the memory needed to store the sparse matrix *Z'* is proportional to $N^{1.5}$ (see fig 4.4): the overall memory occupation of this representation for $\underline{\underline{Z}}$ is proportional to $N^{1.5}$. It is important to note that the impedance matrix $\underline{\underline{Z}}$ is never computed (nor stored) explicitly.

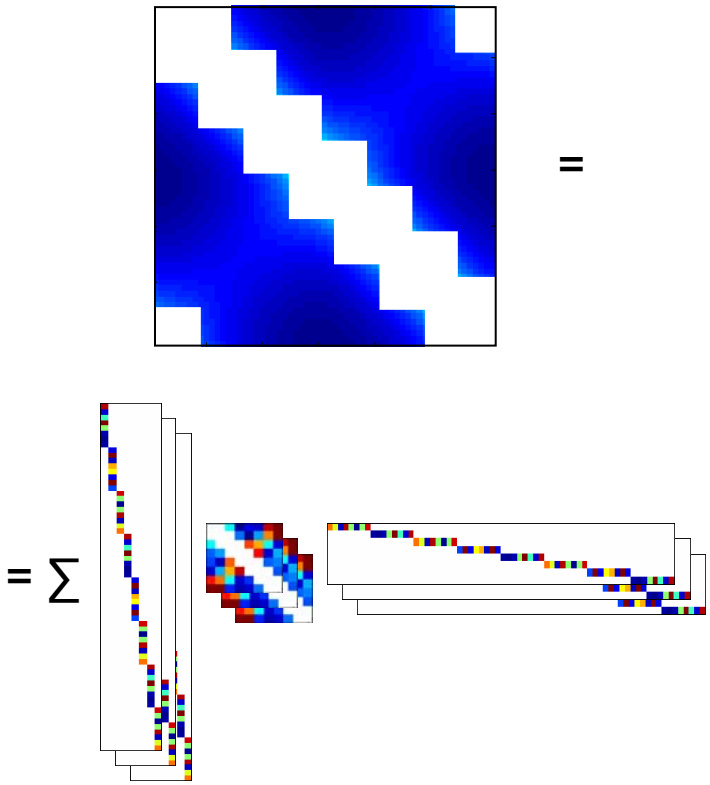


Figure 4.4 – Decomposition of the matrix Z''

From the viewpoint of the computational complexity, the product $\underline{\underline{Z}}\underline{x}$ is carried out with significant savings. The first task concerns the explicit evaluation of the contribution between nearby parts, i.e. the computation of the sparse matrix $\underline{\underline{Z}}$': even if performed in the classical

way, it involves the evaluation of the only nearby interactions, that amounts to a number proportional to $N^{1.5}$. Then, the FMM scheme is applied for taking into account the far interactions. The algorithm involves:

- The aggregation of the source contributions into the clusters
- The translation of the contributions from the source clusters to the observation clusters
- The disaggregation of the contribution from the cluster to the observation points

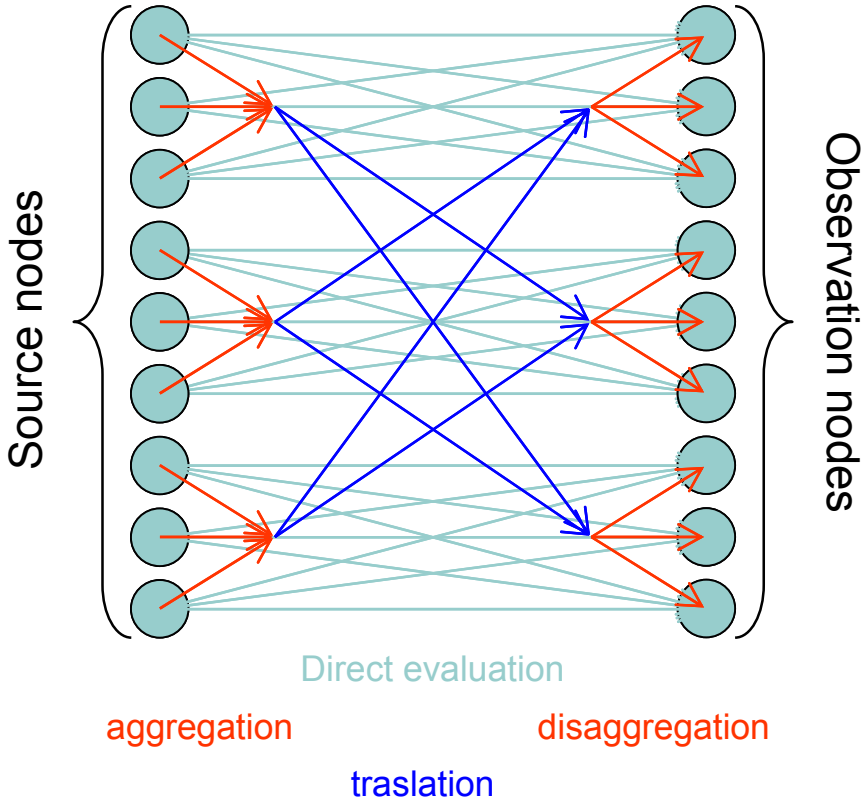


Figure 4.5 – Scheme of the FMM based product. Green links represent the explicitly evaluated nearby interactions, red links the aggregation and disaggregation, blue links the interaction between the multipoles

Each of the three steps is carried out by means of a number of operation proportional to $N^{1.5}$, so that the overall computational complexity of the matrix-by-vector product is $O(N^{1.5})$. The figure 4.5 shows the scheme of the FMM-based multiplication algorithm.
A natural extension of the Fast Multiple Method is the Multi Level Fast Multipole Algorithm [39, 40], where the above described scheme is repeated recursively. The FMM reduces the computational complexity from $O(N^2)$ to $O(N^{3/2})$ by exploiting a two-level scheme. In the same way, further reduction can be obtained by adding more "hub" levels: for example, a three-level scheme leads to a complexity of $O(N^{4/3})$, and so on (see fig. 4.6). It is found that the total memory storage can asymptotically become proportional to $N\log(N)$, while the computational complexity of the matrix by vector multiplication reaches $O(N\log(N))$.

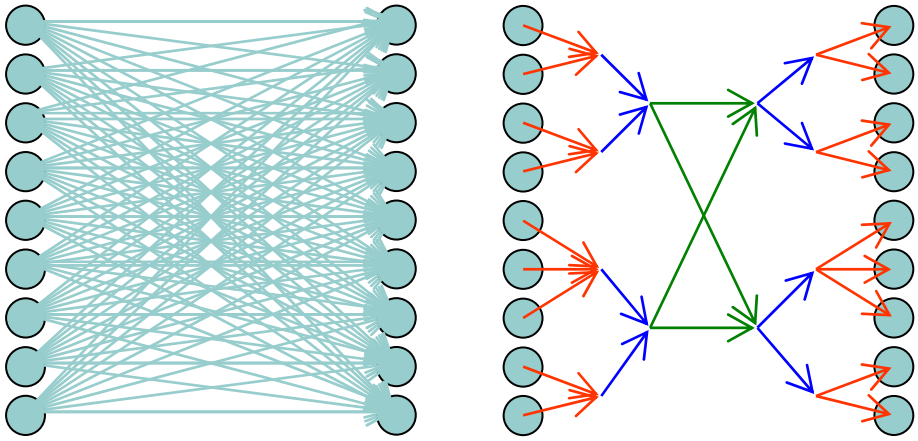


Figure 4.6 – a) Full connected network; b) Network exploiting 2 levels of "hubs" to reduce the number of links

The proposed FMM scheme can be applied to any MoM-like integral formulation of the scattering problem, provided that it is possible to express the far interaction matrix as stated in (4.16).

4.4 The FMM applied to a generic MoM-EFIE formulation

The goal is to find a separable expansion for the Green's function, to obtain the expression (4.15) for the "far interaction" part of the impedance matrix. When dealing with perfect conductive scatterers immersed in free space, in 2D or 3D environment, the Green's function is analytically known in closed form, and the addition theorem can be easily applied.
In the following, we first introduce the general case, with reference to a classical MoM-EFIE formulation of the scattering problem.

4.4.1 Factorization of the interactions

As seen in chapter 3, applying the Method of Moments in the EFIE formulation, the elements of the impedance matrix can be expressed as:

$$Z_{n,n'} = \int_{S_n} \underline{w}_n(\underline{r})\, d\underline{r} \cdot \int_{S_{n'}'} \underline{f}_{n'}(\underline{r}')\underline{\underline{G}}(\underline{r},\underline{r}')\, d\underline{r}' \qquad (4.17)$$

where: $\underline{w}_n$ and $\underline{f}_{n'}$ are the n_th testing and the n'_th expansion functions, defined on the supports S_n and $S'_{n'}$, respectively, and $\underline{\underline{G}}$ is:

$$\underline{\underline{G}}(\underline{r},\underline{r}') = \left(\underline{\underline{I}} + \frac{\nabla\nabla}{k^2}\right) g(\underline{r},\underline{r}') \qquad (4.18)$$

g being the dyadic Green's function:

$$g(\underline{r},\underline{r}') = \frac{\exp(-jk|\underline{r}-\underline{r}'|)}{4\pi|\underline{r}-\underline{r}'|} \qquad (4.19)$$

Let us consider the decomposition of the distance vector into the sum of two vectors (see figure 4.7):

$$\underline{r} - \underline{r}' = \underline{X} + \underline{d} \tag{4.20}$$

The expression of the Green's function becomes:

$$g(\underline{r},\underline{r}') = \frac{\exp(-jk|\underline{X}+\underline{d}|)}{4\pi|\underline{X}+\underline{d}|} \tag{4.21}$$

For $|\underline{X}|>|\underline{d}|$, the above expression can be manipulated by means of the addition theorem for the spherical Bessel functions [61], leading to:

$$\begin{aligned}&\frac{\exp(-jk|\underline{X}+\underline{d}|)}{4\pi|\underline{X}+\underline{d}|} = \\ &= \frac{-jk}{4\pi}\sum_{l=0}^{\infty}(-1)^l(2l+1)j_l(k|\underline{d}|)h_l^{(2)}(kX)P_l(\hat{\underline{d}}\cdot\hat{\underline{X}})\end{aligned} \tag{4.22}$$

where $j_l(\cdot)$ is the spherical Bessel function, $h_l^{(2)}(\cdot)$ is the spherical Hankel function of the second kind, $P_l(\cdot)$ is the Legendre polynomial.

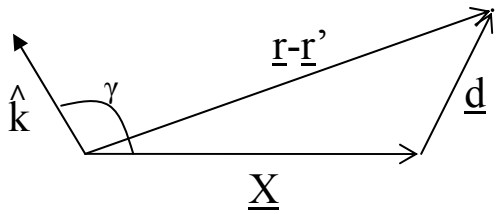


Figure 4. 7 – Decomposition of the distance vector.

The product $j_l P_l$ can be expanded as an integral of propagating plane waves [62]:

$$4\pi j^l(k|\underline{d}|)P_l(\hat{\underline{d}}\cdot\hat{\underline{X}}) = \int d^2\hat{k}\; e^{jk\underline{\hat{k}}\cdot\underline{d}} P_l(\underline{\hat{k}}\cdot\hat{\underline{X}}) \qquad (4.23)$$

where the versor $\hat{k}$ is a generic direction and $\int d^2\hat{k}$ indicates the integral over the unit sphere.
By merging the two results (4.22) and (4.23), and by truncating the infinite sum to $L+1$ terms, we have:

$$\frac{\exp(-jk|\underline{X}+\underline{d}|)}{4\pi|\underline{X}+\underline{d}|} \approx \frac{-jk}{16\pi^2}\int d^2\hat{\underline{k}}\; e^{jk\underline{\hat{k}}\cdot\underline{d}}\cdot \cdot\sum_{l=0}^{L} j^l(2l+1)j_l(k|\underline{d}|)h_l^{(2)}(k|\underline{X}|)P_l(\underline{\hat{k}}\cdot\hat{\underline{X}}) \qquad (4.24)$$

The so obtained expression (4.24) presents factors depending only on $\underline{d}$ or $\underline{X}$. Defining the so called "translation function":

$$T_L(k|\underline{X}|,\gamma) = \frac{-jk}{4\pi}\sum_{l=0}^{L} j^l(2l+1)h_l^{(2)}(k|\underline{X}|)P_l(\cos\gamma) \qquad (4.25)$$

γ being the angle between $\hat{k}$ and $\hat{X}$ (see figure 4.7), the final expression for the Green's function becomes:

$$\frac{\exp(-jk|\underline{X}+\underline{d}|)}{4\pi|\underline{X}+\underline{d}|} \approx \frac{1}{4\pi}\int d^2\hat{\underline{k}}\; e^{jk\underline{\hat{k}}\cdot\underline{d}} T_L(k|\underline{X}|,\underline{\hat{k}}\cdot\hat{\underline{X}}) \qquad (4.26)$$

It is important to note that the translation function doesn't depend on d.
Now, by means of equation (4.17) we want to evaluate the generic element $Z_{n,n'}$ of the impedance matrix, i.e. the interaction between a

generic basis function, say $\underline{f}_{n'}$, and a generic testing function, say $\underline{w}_n$ (see figure 4.8).
As anticipated above, both the basis and testing function sets should be subdivided in localized clusters. If we call M the number of clusters, we have grouped $P=N/M$ functions in each one. Let us label with m the testing clusters and with m' the source clusters. The particular testing or basis function in each cluster is indexed with p and p', respectively. This means that the indexes n and n' have been substituted each by a couple of indexes:

$$n \to (m, p) \quad ; \quad n' \to (m', p') \tag{4.27}$$

Suppose that $\underline{f}_{n'}$ belongs to a cluster centered in $\underline{r}_{m'}$ and the $\underline{w}_n$ belongs to a cluster centered in $\underline{r}_m$. Let us decompose the segment $\underline{r}$-$\underline{r}'$ as follows:

$$\underline{r} - \underline{r}' = \underline{d}_{m,p} + \underline{d}_{m',p'} + \underline{X}_{m,m'} \tag{4.28}$$

where:

$$\underline{d}_1 = \underline{r}_{m'} - \underline{r}' \tag{4.29}$$

$$\underline{d}_2 = \underline{r} - \underline{r}_m \tag{4.30}$$

$$\underline{X} = \underline{r}_c - \underline{r}_{c'} \tag{4.31}$$

$\underline{r}_c$ and $\underline{r}_{c'}$ being the positions of the test and of the source clusters, respectively.

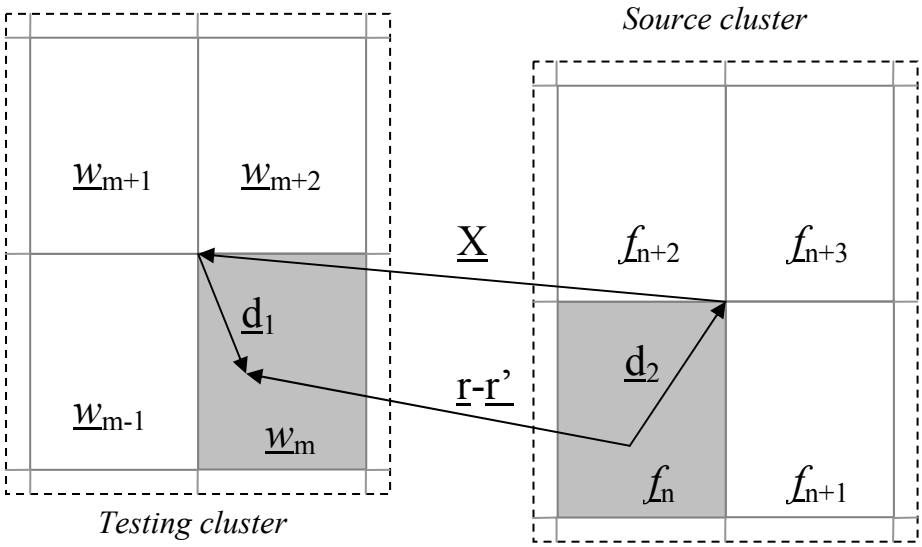


Figure 4.8 – Decomposition of the segment $\underline{r}_1$-$\underline{r}_2$

After these considerations, assuming $\underline{d} = \underline{d}_1 + \underline{d}_2$, the expression (4.26) can be rewritten:

$$\frac{\exp(-jk\left|\underline{X} + \underline{d}_1 + \underline{d}_2\right|)}{4\pi\left|\underline{X} + \underline{d}_1 + \underline{d}_2\right|} \approx$$
$$\approx \frac{1}{4\pi}\int d^2\hat{\underline{k}}\, e^{jk\underline{\cdot}\underline{d}_1} e^{jk\underline{\cdot}\underline{d}_2} T_L(k\left|\underline{X}\right|, \hat{\underline{k}}\cdot\hat{\underline{X}}) \tag{4.32}$$

and the $\underline{\underline{G}}$ function:

$$\underline{\underline{G}}(\underline{r},\underline{r}') \approx$$
$$\approx \frac{1}{4\pi}\left(\underline{\underline{I}} + \nabla\nabla\right)\int d^2\hat{\underline{k}}\, e^{jk\underline{\cdot}\underline{d}_1} e^{jk\underline{\cdot}\underline{d}_2} T_L(k\left|\underline{X}\right|, \hat{\underline{k}}\cdot\hat{\underline{X}}) \tag{4.33}$$

so that, from eq. (4.17), the expression for the generic element of the impedance matrix is:

$$Z_{n,n'} = \int_{S_n} \underline{w}_n(\underline{r})\, d\underline{r} \cdot \int_{S'_{n'}} \underline{f}_{n'}(\underline{r}') \frac{1}{4\pi} \left(\underline{\underline{I}} + \nabla\nabla\right) \cdot \\ \cdot \int d^2\hat{\underline{k}}\, e^{j\underline{k}\cdot\underline{d}_1} e^{j\underline{k}\cdot\underline{d}_2} T_L(k|\underline{X}|, \hat{\underline{k}}\cdot\hat{\underline{X}})\, d\underline{r}' \tag{4.34}$$

In the integration domain $S_n \times S'_{n'}$, $\underline{r}$ depends only on $\underline{d}_2$ and $\underline{r}'$ depends only on $\underline{d}_1$: as a consequence, the $\underline{\underline{G}}$ is a function of the couple $(\underline{d}_1, \underline{d}_2)$, and does not vary with $\underline{X}$. As a consequence, the $\nabla\nabla$ operator is applied only on the exponential terms $\exp(j\underline{k}\cdot\underline{d}_{1,2})$ and it can be substituted by the operator $-\hat{\underline{k}}\hat{\underline{k}}$:

$$\underline{\underline{Z}}_{n,n'} = \frac{1}{4\pi} \int d^2\hat{\underline{k}} \int_{S_n} d\underline{r} \left(\underline{\underline{I}} - \hat{\underline{k}}\hat{\underline{k}}\right) \exp(j\underline{k}\cdot\underline{d}_2) \underline{w}_n(\underline{r}) \cdot \\ \cdot \int_{S'_{n'}} d\underline{r}' \left(\underline{\underline{I}} - \hat{\underline{k}}\hat{\underline{k}}\right) \exp(j\underline{k}\cdot\underline{d}_1) \underline{f}_{n'}(\underline{r}') T_L(k|\underline{X}|, \hat{\underline{k}}\cdot\hat{\underline{X}}) \tag{4.35}$$

where we have inverted the integration order. We can define the "excitation vectors" $\underline{V}_{m,p}$ and $\underline{V'}_{m',p'}$:

$$\underline{V}'_{m',p'}(\hat{\underline{k}}) = \int_{S'_{n'}} \left(\underline{\underline{I}} - \hat{\underline{k}}\hat{\underline{k}}\right) \exp(j\underline{k}\cdot\underline{d}_1) \underline{f}_{n'}(\underline{r}') d\underline{r}' \tag{4.36}$$

$$\underline{V}_{m,p}(\hat{\underline{k}}) = \int_{S_n} \left(\underline{\underline{I}} - \hat{\underline{k}}\hat{\underline{k}}\right) \exp(j\underline{k}\cdot\underline{d}_2) \underline{w}_n(\underline{r}) d\underline{r} \tag{4.37}$$

Then we have:

$$\underline{\underline{Z}}_{n,n'} = \frac{1}{4\pi}\int d^2\hat{\underline{k}}\,\underline{\underline{V}}_{m,p}(\hat{\underline{k}})T_L(k|\underline{X}|,\hat{\underline{k}}\cdot\hat{\underline{X}})\underline{\underline{V}}'_{m',p'}(\hat{\underline{k}}) \quad (4.38)$$

The integral over the unit sphere $\int d^2\hat{\underline{k}}$ can be discretized by means of a summation of contributes evaluated on a suitable set of N_K discrete directions $\hat{\underline{k}}_1 \ldots \hat{\underline{k}}_{N_K}$:

$$\underline{\underline{Z}}_{n,n'} = \\ = \frac{1}{4\pi}\sum_{i=1}^{N_K}\Delta\hat{\underline{k}}\,\underline{\underline{V}}_{m,p}(\hat{\underline{k}}_i)T_{L,m,m'}(k|\underline{X}|,\hat{\underline{k}}_i\cdot\hat{\underline{X}})\underline{\underline{V}}'_{m',p'}(\hat{\underline{k}}_i) \quad (4.39)$$

The (4.40) is a factorized expression for the impedance matrix elements according to the scheme (4.16) described in the previous paragraph. As can be seen, the dependence of the matrix element on the distance vector has been disaggregated.
Note that the two indexes *m* and *m'* have been added as pedices to the translation function T_L, since it depends on the particular pair of considered clusters.
Relations (4.36) and (4.37) shows that the excitation vectors $\underline{\underline{V}}'_{m',p'}(\hat{\underline{k}}_i)$ and $\underline{\underline{V}}_{m,p}(\hat{\underline{k}}_i)$ can be evaluated once (in the discrete set of values $\hat{\underline{k}}_1 \ldots \hat{\underline{k}}_K$) for each basis and testing function, and plays the roles of aggregation and disaggregation functions, respectively. On the other hand, the translation functions $T_{L,m,m'}(k|\underline{X}|,\hat{\underline{k}}_i\cdot\hat{\underline{X}})$ can be computed once for each couple of clusters *m* and *m'*.
In the FMM implementation, not all the possible couples of cluster have to be actually considered: in fact, the factorization for the Green's function is valid under the condition $|\underline{d}|<|\underline{X}|$. As a consequence, the m-t testing cluster and the *m*'-th source cluster are considered "nearby" if their distance $|\underline{X}_{m,m'}|$ is larger than the radius of the cluster. This forces to exclude from the fast scheme the evaluation of the interactions between adjacent clusters, that are instead taken into account in the near interaction matrix *Z*', and directly computed.

In order to estimate the computational complexity related to the FMM-based matrix-by-vector product, it is necessary discuss validity of the approximation previously introduced.

4.4.2 Truncation and discretization error

The first approximation we made concerns the truncation to L+1 terms of the summation over the index l in the relation (4.22). It is well known that the values of the Bessel's functions $j_l(x)$ and $h_l(x)$ stays bounded until the order is lower than the argument. When l exceeds the value of x, $h_l(x)$ starts to grow, and $j_l(x)$ decays rapidly. We can reasonably suppose that for $l>|kd|$ the factor $j_l(kd)$ becomes almost zero, so that the contributions that the further terms give to the summation become negligible. Anyway, when the index l exceeds $|kX|$, because of the growth of the Hankel's function $h_l(kX)$, the integrand starts to oscillate: due to the finite machine precision, the summation of additional terms should only add numerical error to the solution. In determining L, a semi-empirical formula has been proposed by Coifman, Rokhlin and Wandzura [63]:

$$L\ (kd_{\max}) = kd_{\max} + \alpha \ln(kd_{\max} + \pi) \qquad (4.40)$$

where $d_{\max}$ is the maximum distance between a generic point in any basis/testing function and the relative cluster center, and α is a parameter that depends on the prescribed precision.
The second approximation concerns the discretization of the integral over the unit sphere into a finite summation made in (4.39). The number N_K of directions $\hat{\underline{k}}_i$ must be enough to ensure accuracy for all spherical harmonics of order $l<L$. Being $\hat{\underline{k}}_i = (1, \theta_i, \varphi_i)$, a simple method [63] consider the couples (θ_i, φ_i) wherein θ_i corresponds to the zeros of the Legendre polynomial, $P_L(\cos\theta_i)$, and the azimutal angles φ_i are $2L$ equally spaced points [62]. Thus, the total number of

directions would be $N_K=2L^2$, which is essentially proportional to $\underline{d}_{\max}^2$.

4.4.3 Evaluation of the computational costs

It is now possible to assess the cost savings, in terms of memory occupation and computational complexity, due to the introduction of the FMM.
Once the splitting of the impedance matrix into a "near" and a "far" contribution (eq. (4.13)) has been obtained , the matrix-by-vector multiplication required by the i_th step of the iterative solution process can be written as:

$$\underline{E}^{(i)} = \underline{\underline{Z}} \cdot \underline{I}^{(i)} = \underline{\underline{Z}}' \cdot \underline{I}^{(i)} + \underline{\underline{Z}}'' \cdot \underline{I}^{(i)} \qquad (4.41)$$

where $I^{(i)}$ contains the coefficients representing the i-th estimate of the current distribution and $\underline{E}^{(i)}$ the coefficients of the radiated field.
The matrix *Z*', describing the direct interactions between the basis and testing functions corresponding to adjacent clusters, is largely sparse. Since any cluster contains *N/M* functions, the total number of non-zero elements of the matrix *Z*' is proportional to N^2/M. As a consequence, the multiplication with the vector I_i requires a computational complexity O(N^2/M).
The matrix Z'', which accounts for the FMM scheme of interactions, is never computed explicitly. The role of the impedance matrix is played by the excitation vectors $\underline{V}'_{m',p'}(\hat{\underline{k}}_i)$, $\underline{V}_{m,p}(\hat{\underline{k}}_i)$ and the translation functions $T_{L,m,m'}(k|\underline{X}|, \hat{\underline{k}}_i \cdot \hat{\underline{X}})$.
The key parameter of interest is the number of directions $N_K \propto \underline{d}_{\max}^2$, that it proportional to the cluster area.
Once the spatial distribution of the basis functions is chosen, N_K is proportional to the number of functions in each cluster, i.e., to *N/M*.

For any of the N basis and any of the N testing functions, the excitation vectors should be evaluated in N_K directions and N^2/M operations are involved. Obviously, the storage of the results require a memory occupancy proportional to N^2/M .

For any of couple of non-adjacent clusters, the functions $T_{L,m,m'}(k|\underline{X}|, \underline{\hat{k}}_i \cdot \underline{\hat{X}})$ should be evaluated in the N_K directions. Since the number of couples of clusters is essentially proportional to M^2, this task requires $M^2*N/M=MN$ operations, and a memory storage also proportional to MN.

The pre-computation of both the near interaction matrix, the excitation vectors and the translation functions is equivalent to the evaluation of the complete MoM impedance matrix.

Accordingly, the memory storage requirement has been reduced from N^2 elements to a value of order $O(N/M + MN)$. The computational effort also reduces from $O(N^2)$ to $O(N/M) + O(MN)$, even if from the computational point of view the most important role is played by the multiplication (4.41) required by the iterative solvers, rather that by the FMM pre-computation.

The matrix-by-vector multiplication will follow the scheme proposed in figure 4. At first, we rewrite the (4.41) in terms of the pre-computed data. For the generic n_th component of the vector $\underline{E}^{(i)}$:

$$E_n{}^{(i)} = \sum_{n'=1}^{N} Z'_{n,n'} \cdot I_{n'}{}^{(i)} + \sum_{n'=1}^{N} Z''_{n,n'} \cdot I_{n'}{}^{(i)} \qquad (4.42)$$

Both Z’ and Z’’ are sparse matrices. In fact, if the n_th testing function and the n’_th basis function are in adjacent clusters, the second summation in (4.42) is zero, and the n’th contribution to En just comes from the first summation. Elsewhere, when evaluating the interaction between far away functions, only the second summation is different from zero. Defining with Q_n the set of the indexes of the basis function adjacent to the n_th test function, we can rewrite the (4.42):

$$E_n^{(i)} = \sum_{n' \in Q_n} Z'_{n,n'} \cdot I_{n'}^{(i)} + \sum_{n' \notin Q_n} Z''_{n,n'} \cdot I_{n'}^{(i)} \tag{4.43}$$

Since the number of elements in Q_n is proportional to M, the evaluation of the first summation, for all the N components of $\underline{E}^{(i)}$, involves MN operations.
The second contribution in (4.43) can be rewritten in terms of translation and excitation vectors. In fact, using the correspondence between indexes seen in (4.27), it can be expressed as:

$$\begin{aligned} \sum_{n' \notin Q_n} \underline{\underline{Z}}''_{n,n'} \cdot I_{n'}^{(i)} &= \frac{1}{4\pi} \sum_{i=1}^{N_K} \Delta\hat{\underline{k}} \, \underline{V}_{m,p}(\hat{\underline{k}}_i) \cdot \\ \cdot \sum_{m' \notin R_n} T_{L,m,m'}(k|\underline{X}|, \hat{\underline{k}}_i \cdot \hat{\underline{X}}) & \sum_{p' \in P_{m'}} \underline{V}'_{m',p'}(\hat{\underline{k}}_i) I_{n'}^{(i)} \end{aligned} \tag{4.44}$$

where R_n contains the indexes of the clusters adjacent to the n_th testing function, and $P_{m'}$ the set of the sub-indexes inside the m'_th cluster.
An algorithm consisting of three steps allows to evaluate the expression (4.44) for all the values of $n \in \{1,..N\}$.

- *1° step: aggregation.* The expression $\underline{V}'_{m',p'}(\hat{\underline{k}}_i) I_{n'}^{(i)}$ is evaluated for any value of n and in any of the N_K directions $\hat{k}_i$, involving a number of operations proportional to $N*N/M=N^2/M$. The values, that represent the contribution of the whole source cluster to the radiation, are aggregated in the quantities:

$$s_{m'}(\hat{\underline{k}}_i) = \sum_{p' \in P_{m'}} \underline{V}'_{m',p'}(\hat{\underline{k}}_i) I_{n'}^{(i)} \tag{4.45}$$

- *2° step: translation.* The expression:

$$g_m(\hat{\underline{k}}_i) = \sum_{m' \notin R_n} T_{L,m,m'}(k|\underline{X}|, \hat{\underline{k}}_i \cdot \hat{\underline{X}}) s_{m'}(\hat{\underline{k}}_i) \qquad (4.46)$$

is evaluated for any of the M testing clusters, in any of the N_K directions $\hat{k}_i$. The computational effort is proportional to $N_K M^2 = NM$.

- *3° step: disaggregation.* The expression $\underline{V}_{m,p}(\hat{\underline{k}}_i) g_m(\hat{\underline{k}}_i)$, representing the contributions of the field radiated from the far sources into the n_th testing function, is evaluated for any n (i.e., for any couple of m and p) and in any direction $\hat{k}_i$, then the sum over i in the (4.44) is performed. In this way, all the N components of the vector $\underline{E}^{(i)}$ are evaluated by means of $NN_K = N^2/M$ operations.

The above described aggregation-translation-disaggregation scheme is introduced in the iterative process of solution. Accordingly, the overall complexity of the solution process can be expressed as:

$$O(N_{iter} NM) + O(N_{iter} N^2/M) \qquad (4.47)$$

where N_{iter} is the number of iteration needed by the method to converge.

The choice of M should ensure the minimum computational effort: it is clear from (4.47) that the optimal choice is $M = \sqrt{N}$. Taking into account this consideration, and supposing that the number N_{iter} is independent from N, the computational complexity of the FMM-based solution process becomes $O(N^{1.5})$. This choice of M also minimizes the memory needed to store the near interaction and the translation and excitation vectors, that also becomes proportional to $N^{1.5}$.

4.5 Application of the FMM to the Method of Auxiliary Sources in 2D

The FMM is a very powerful technique and can be successfully applied to speed up the solution process of the linear systems coming

from a MoM (or other IF-based) implementation that utilize localized basis and testing functions.

In this paragraph we show how to apply the FMM paradigm to the Method of Auxiliary Sources in 2D. The procedure essentially resembles that of the MoM case.

The first problem concerns the factorization of the Green function. In a 2D case, the ASs are filamentary currents, whose field behaves like a Hankel function of second kind $H_0^{(2)}(kr)$, $\underline{r}$ being the distance between the source and the observation point and k the propagation constant. Putting $\underline{r}=\underline{X}+\underline{d}$, for $|\underline{d}|<|\underline{X}|$ we can apply the addition formula [61]:

$$H_0^{(2)}(kr) = \sum_{l=-\infty}^{+\infty} H_l^{(2)}(kX) J_l(kd)\cos(l\alpha) \qquad (4.48)$$

J_l being the Bessel function of first kind and α the angle between $\underline{X}$ and $\underline{d}$. Since the Bessel functions decay rapidly for $|l|>k|\underline{d}|$, the infinite summation may be truncated between $-L$ and $+L$, with $L> k|\underline{d}|$.

Moreover, we have [63]:

$$J_l\left(k|\underline{d}|\right)\cos\left(l\alpha\right) = \frac{j^{-l}}{2\pi}\int_{-\pi}^{\pi} e^{jk|\underline{d}|\cos(\varphi)}\cos\left(l(\varphi-\alpha)\right)d\varphi \qquad (4.49)$$

From (4.48) and (4.49) we find:

$$H_0^{(2)}(k|\underline{X}+\underline{d}|) \cong \int d\hat{\underline{k}}\, e^{j\underline{k}\cdot\underline{d}} \sum_{l=-L}^{+L} \frac{j^{-l}}{2\pi} H_l^{(2)}(k|\underline{X}|)\cos(l\gamma) \qquad (4.50)$$

where $\int d\hat{\underline{k}}$ indicates the integral over the unit circle and γ is the angle between the vectors $\underline{X}$ and $\hat{\underline{k}}$.

Let N be the number of auxiliary sources and testing points (they play the roles of the basis and testing functions, respectively). As seen in the previous section, both the sets can be subdivided into M localized groups, so that each cluster contains $P=N/M$ elements. Let us label

with *m* the testing point clusters and with *m'* the source clusters. The particular testing point/source point in each cluster is indexed with *p* and *p'*, respectively. In this way, the vector joining the source and the testing point can be written as $\underline{r} = \underline{X}_{m,m'} + \underline{d}$, where $\underline{d} = \underline{d}_{m,p} - \underline{d}_{m',p'}$, $\underline{X}_{m,m'}$ being the vector joining the centres of the clusters, $\underline{d}_{m,p}$ and $\underline{d}_{m',p'}$ being those from the centre of the cluster to the testing/source point, respectively (see figure 4.9).

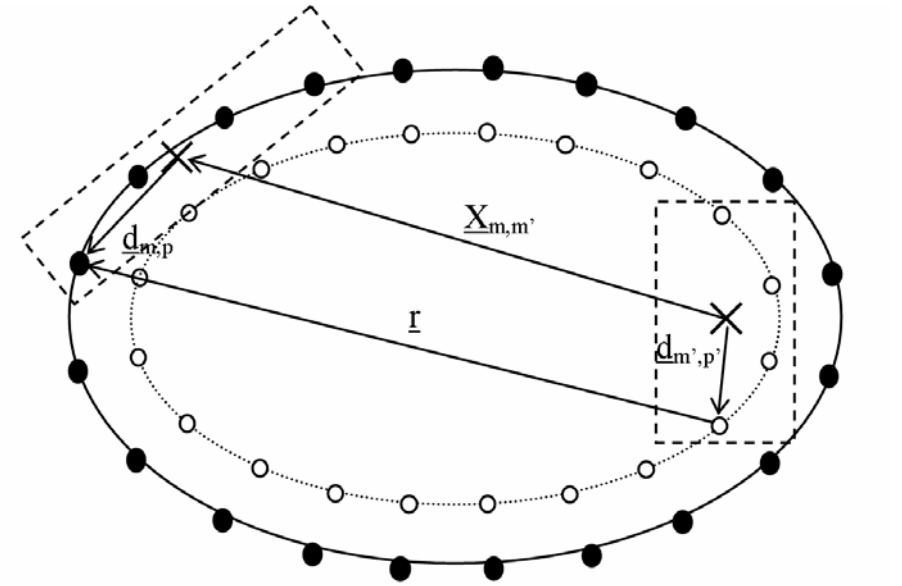


Figure 4.9 - Decomposition of the distance $\underline{r} = \underline{X}_{m,m'} + \underline{d}_{m,p} - \underline{d}_{m',p'}$. *Black circles denotes the testing points, white circles the auxiliary sources, dotted curves delimit the clusters interested in interaction, crosses indicates the cluster centres*

Now we want to consider the interaction between a generic source point (*m',p'*) and a generic testing point (*m,p*). The integral in (4.50) can be numerically evaluated over a discrete set of N_k directions $\hat{\underline{k}}_i$:

$$H_0^{(2)}(k|\underline{X}+\underline{d}|)=\sum_{\hat{k}_i=1}^{N_k} V_{m,p}(\hat{\underline{k}}_i)V_{m',p'}(\hat{\underline{k}}_i)\Delta\hat{\underline{k}}T_{m,m'}(\hat{\underline{k}}_i) \qquad (4.51)$$

being:

$$T_{m,m'}(\hat{\underline{k}}_i)=\sum_{l=-L}^{+L}\frac{j^{-l}}{2\pi}H_l^{(2)}(k|\underline{X}_{m,m'}|)\cos(l\gamma) \qquad (4.52)$$

$$V_{m,p}(\hat{\underline{k}}_i)=e^{jk\hat{\underline{k}}_i\cdot\underline{d}_{m,p}} \quad ; \quad V_{m',p'}(\hat{\underline{k}}_i)=e^{-jk\hat{\underline{k}}_i\cdot\underline{d}_{m',p'}} \qquad (4.53)$$

The relation (4.53) requires $|\underline{d}|<|\underline{X}|$, so it can be used to evaluate the elements of the MAS impedance matrix which represent the interaction between non-adjacent clusters. The field radiated at the n-th testing point by the non-adjacent source clusters, i.e. the n-th component of the product $\underline{\underline{Z}}''\underline{I}$, can be factorized as:

$$E_{m,p} = \\ =\sum_{\hat{k}_i=1}^{N_k} V_{m,p}(\hat{\underline{k}}_i)\Delta\hat{k}\sum_{m'=1}^{M}T_{m,m'}(\hat{\underline{k}}_i)\sum_{p'=1}^{P}V_{m',p'}(\hat{\underline{k}}_i)I_{m',p'} \qquad (4.54)$$

As seen, the equation (4.54) represents an aggregation-translation-disaggregation scheme for the (impedance) matrix-by-(source) vector multiplication similar to the one considered in the previous section. Being L and N_K essentially proportional to the radius of the clusters, as seen before, the application of the FMM scheme to a 2D MAS implementation allows to reduce the memory storage from N^2 to $N^{1.5}$, and the computational complexity of the matrix-by-vector multiplication from $O(N^2)$ to $O(N^{1.5})$.
The FMM-MAS technique has been applied to the analysis of large PEC scattering objects, achieving significant results in terms of reduction of computational costs. Some interesting numerical results can be found in chapter 6.

Chapter 5

A non-redundant implementation of the Method of the Auxiliary Sources

In order to take full advantage of the ability of the MAS in reducing the number of unknowns, we want to exploit the results of the theory of the optimal sampling of the scattered fields [3, 50, 51], providing reliable guidelines to the positioning of both the testing points and the auxiliary sources. According to this theory , the field radiated by a scattering or radiating system over a smooth curve, not closer than some wavelength from the radiators, can be approximated by means of a band-limited function of the curve parameterization with a bandwidth w. The approximation error is a step-like function of w and decreases drastically for values of w, hence of the spatial frequency, higher than β [50]. Based on this basic result, the criteria underlying the estimation of the number of the degrees of freedom of the field radiated by a given set of scatterers or radiators are given in [2]. In particular, avoiding the case of superdirective sources, it is shown that they can be estimated by means of only geometrical considerations, being proportional to the extension of the surface of the minimum convex envelope including the scattering/radiating system. It is also shown that the field can be represented with the desired accuracy by considering a number of sampling points slightly larger than the degrees of freedom of the field.

In common MAS implementations, usually the auxiliary surface is chosen conformal to the boundary, and the testing points are usually placed uniformly along the curvilinear abscissa. However, the application of some results deriving from the optimal sampling theory suggest different choices that significantly improve the performances of the Method of Auxiliary Sources applied to extended PEC objects, and allows the use of a minimum set of unknowns.

5.1 Spatial bandwidth and optimal sampling of the scattered fields

In this paragraph, a brief description of the theory on the spatial bandwidth and the optimal sampling of the scattered field is reported.

5.1.1 Problem statement

Let us consider a set of scattering objects in free space, and denote with *a* the radius of the smallest sphere *B* enclosing all of them (see Figure 1). In the following, an *exp*(*jωt*) time dependence is supposed and suppressed. The field scattered at the point $\underline{r}$, belonging to a given observation domain *M*, according to [50], can be expressed as:

$$\underline{E}_0(\underline{r}) = \int_D \underline{\underline{G}}_0(\underline{r},\underline{r}') \cdot \underline{J}(r')d\underline{r}' \tag{5.1}$$

where $\underline{\underline{G}}_0$ is the dyadic Green's function and $\underline{J}$ are the (or induced) surface current on the domain *D*, given by the surfaces of the scattering objects included in *B* and $\underline{r}$' is the position vector of a point of *D*. In the free space we have:

$$\underline{\underline{G}}_0(\underline{r},\underline{r}') = -\frac{j\omega\mu}{4\pi}\left[\bar{I} + \frac{\nabla\nabla}{\beta^2}\right]\frac{\exp(-j\beta R)}{R} \tag{5.2}$$

where $R=|\underline{r}-\underline{r}'|$ and $\beta = \omega\sqrt{\varepsilon\mu}$.

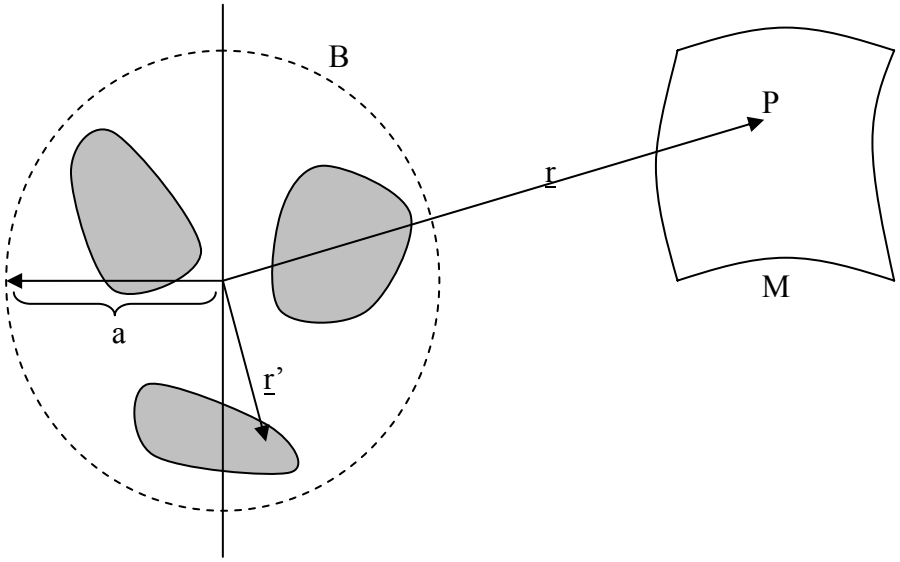


Figure 5.1 – Geometry of the problem.

The observation point $\underline{r}$ can be placed in the complex extension of the domain *M*. The only singularity for the function (5.2) occurs when the source point $\underline{r}'$ coincides with the observation point $\underline{r}$, i.e. when R=0. Such situation is easily avoided when the observation domain is placed away from the ball *B*.
The exponential factor contained in (5.2), when *r* is allowed to take complex values, can grow indefinitely at infinity. This factor can be conveniently singled out by considering the *reduced field*:

$$\underline{E}(\underline{r}) = \underline{E}_0(\underline{r})\exp(j\beta r) \qquad (5.3)$$

In this way, a "reduced" Green's function can be defined:

$$\underline{\underline{G}}(\underline{r},\underline{r}') = -\frac{j\omega\mu}{4\pi R}\exp(j\beta(r-R))\underline{\underline{N}}(\underline{r},\underline{r}') \qquad (5.4)$$

where the matrix $\underline{\underline{N}}(\underline{r},\underline{r}')$ is defined as:

$$\underline{\underline{N}}(\underline{r},\underline{r}') = \left[\bar{I} + R\exp(j\beta R)\frac{\nabla\nabla}{\beta^2}\frac{\exp(-j\beta R)}{R} \right] \tag{5.5}$$

and the field expression becomes:

$$\underline{E}(\underline{r}) = \int_D \underline{\underline{G}}(\underline{r},\underline{r}')\cdot\underline{J}(r')d\underline{r}' \tag{5.6}$$

It is interesting to note that the norm of the matrix (5.5) approaches unity as βR becomes large. Moreover, the exponential term grows with (r-R), that stays bounded when R goes to infinite. As a consequence, we can say that any scattered field is an analytic function, regular in a complex extension of the real observation domain M [50].

5.1.2 Approximating the field with a band-limited function

As shown in [50], these results allow to approximate the radiated fields by means of bandlimited functions. As an instance of observation domain, let us consider an indefinite analytical curve S, external to the ball B. To avoid any spurious singularities due to the properties of the coordinate mapping, the arc length s, normalized by the minimum distance between the curve S and the origin O is taken as the curve parameter, s being the curvilinear coordinate on S. Being $\underline{r}=\underline{r}(s)$, also the Green's function (5.4) can be rewritten as a function of the normalized arclength:

$$\underline{\underline{G}}(s,\underline{r}') = -\frac{j\omega\mu}{4\pi R}\exp(j\psi(s,\underline{r}'))\underline{\underline{N}}(s,\underline{r}') \tag{5.7}$$

with

$$\psi(s,\underline{r}') = \beta\left(\underline{r}(s) - R(s,\underline{r}')\right) \tag{5.8}$$

The best approximation of the field *E(s)* with a band-limited function is the projection of the field on the set of all the functions defined in the observation domain and band-limited to *w*. By applying such projection to the expression (5.6), the evaluation of the approximation error is transferred from the scattered field to the Green's function. We obtain that the maximum error can be estimated with [50]:

$$\left|\Delta \underline{E}\right| \le \max_{s,\underline{r}'}\left[\left|\Delta \underline{\underline{G}}(s,\underline{r}')\right|\right]\cdot \int_D \left|\underline{J}(r')\right| d\underline{r}' \qquad (5.10)$$

where $\Delta \underline{E}$ is the difference field between $\underline{E}(s)$ and its band-limited version, and $\Delta \underline{\underline{G}}$ is the difference between the Green's function and its band-limited version.
In this way the problem turns into the evaluation of the error Green's function. To this purpose, a modified integral representation is exploited to perform an asymptotical evaluation. It is found that, for a given source point at $\underline{r}'$, the value of $\Delta \underline{\underline{G}}$ is comparable with $\underline{\underline{G}}$ if do exist stationary points defined by the condition:

$$w(\underline{r}') = \pm \frac{\partial}{\partial \xi}\psi(s(\xi),\underline{r}') \qquad (5.11)$$

where ξ is an arbitrary, analytic curve parameterization. the normalized arc length. If we define:

$$w_0(\underline{r}') = \max_{\xi}\left|\frac{\partial}{\partial \xi}\psi(s(\xi),\underline{r}')\right| \qquad (5.12)$$

we have that no stationary points arise for $w > w_0$, and $\Delta \underline{\underline{G}}$ is expected to be very small [50]. It is interesting to study the behavior of the error in the region where it is small, i. e. for w sufficiently larger than w_0, where it is found that:

$$\max_{s}\left[\left|\Delta\underline{\underline{G}}(s,\underline{r}')\right|\right]\approx A\frac{\exp\left(-\frac{3}{2}\sigma^{3/2}\right)}{\sqrt{2\pi\sigma^{3/2}}} \tag{5.13}$$

A being a constant, and

$$\sigma(\underline{r}')=\left(2\Big/\left|\frac{\partial^3}{\partial\xi^3}\psi(\xi,\underline{r}')\right|_{\xi=\xi_0}^{1/3}(w-w_0)\right) \tag{5.14}$$

ξ_0 being the point of maximum for $\left|\frac{\partial}{\partial\xi}\psi(s(\xi))\right|$.

The evaluation of the band-limitation error requires the maximization of the expression (5.13) with respect to $\underline{r}'$. Aiming to consider any possible source location inside the envelope, we must enforce the condition:

$$w > W = \max_{\underline{r}'} w_0(\underline{r}') \tag{5.15}$$

in order to avoid large errors. The maximization of the (5.13) and its substitution into the (5.10) gives the relation:

$$|\Delta\underline{E}|\le\frac{\omega\mu}{4\pi d}\left[\int_D|\underline{J}(r')|d\underline{r}'\right]\frac{\exp\left(-\frac{2}{3}\sigma_m^{3/2}\right)}{\sqrt{2\pi\sigma_m^{3/2}}} \tag{5.16}$$

with $\sigma_m=\min_{\underline{r}'}\sigma(\underline{r}')$.

5.1.3 Estimation of the spatial bandwidth

The band-limitation error expressed in (5.16), shows a dependence from the entity of the induced currents, that are obviously finite, and from a factor related to the "band excess", w-W. In fact, it is clear that the value of σ_m is controlled by the choice of the band w and by the value of W, since it is always w-w_0>w-W. The maximum bandwidth can be evaluated by means of geometrical considerations.
In particular, an upper bound for W is found to be always lower or equal to $\sqrt{2}\beta a$, and we have $W \approx \beta a$ as the distance between the scattering envelope and the observation curve becomes slightly larger than a. The quantity:

$$\chi = \frac{w}{W} \approx \frac{w}{\beta a} \tag{5.17}$$

is defined as the excess bandwidth factor, an we can express:

$$w - W \approx \beta a(\chi - 1) \tag{5.18}$$

Since the value of σ_m is controlled by the excess bandwidth, as said above, it is clear that the error (5.16) decays exponentially as w exceeds the value βa, showing a step-like behavior. Thus, it is possible to identify βa as the effective bandwidth of the scattered field.
The possibility of approximating the fields over arbitrary observation domains by means of band-limited functions has been successfully exploited to propose an innovative approach to the sampling problem [51]. With reference to an one-dimensional observation domain (i.e. a generic curve), starting from only geometrical information about the scattering system, it is possible to evaluate the maximum field bandwidth and, as a consequence, the minimum sampling rate needed to perform a spatial sampling with an acceptable error.
For a closed curve, the number of samples needed to express any scattered field is finite, and can be seen that it is proportional to βa. In particular, it is found that the minimum number of samples needed to

reconstruct the field over a sphere is finite, approximately $4(\beta a)^2/\pi$, and practically equal to the degrees of freedom of the field, i.e. the number of relevant spherical harmonics $(\beta a)^2$.

For an unbounded line (see fig. 5.2), the sampling of the field at a fixed rate, related to the estimated effective bandwidth, provides an infinite number of samples. Since the field on any point outside the sphere can be obtained by the tangential components on its surface, the degrees of freedom are finite in the whole space, and it is desirable to have a technique able to represent the field with a finite, non redundant number of samples, even on an unbounded domain.

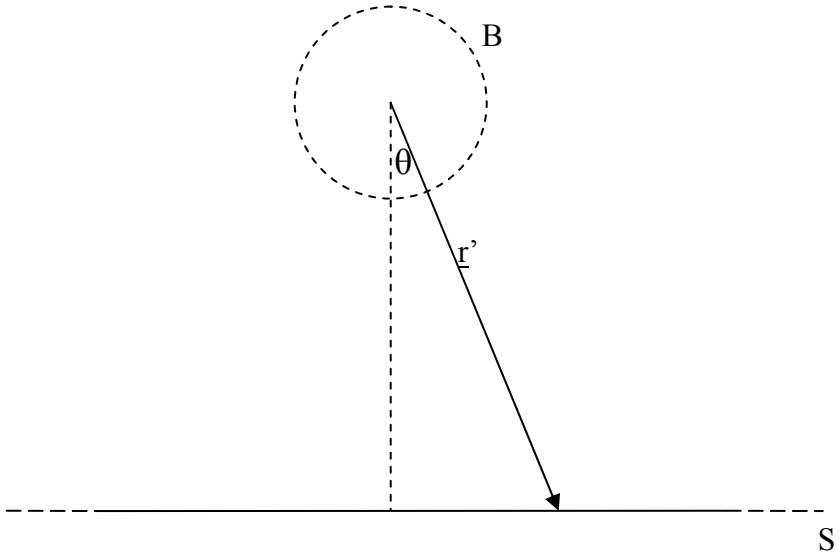


Figure 5.2 – Example of unbounded observation domain

5.1.4 Introducing the generalized field reduction

A non-redundant sampling approach to the above issue has been presented in 1994 [64, 65, 66]. It relies on the generalization of the field's reduction; let us consider:

$$F(\xi) = \exp(j\gamma(\xi))E(\xi) \qquad (5.19)$$

where $\gamma(\xi)$ is a the generalized reduction function, and $\xi(s)$ a generic parameterization of the curve, s indicating the arc length as usual. The aim is to determine both those functions in order to minimize the number of samples needed. In fact, it is clear that the local bandwidth can vary with the value of ξ, and this implies that the value of W can be larger that the effective local bandwidth in most of the observation domain. To avoid this problem, the value of w_0 should be constant with ξ, and, obviously, equal to W. By expressing w_0 for the generalized reduced field, it is possible to determine the expression of both the reduction function and the curve parameterization matching this condition. For example, with reference to the previously considered case of sources enclosed in a circle (see fig. 5.2), the curve parameterization is found to coincide with the angular variable θ. In this way, it turns out that the sampling at a constant rate (with respect to the parameter) involves a finite number of sampling points, even on the considered unbounded observation domain. In a certain way, we can imagine that the degrees of freedom of the field "move along the radii" of a spherical reference system centered in the center of the sphere enclosing the sources/scatterers.

5.1.5 Extension to non-spherical scattering envelopes

In many cases of practical interest, for instance when dealing with planar antennas, the minimum spherical envelope is not the most realistic way to take into account the scattering geometry. Important achievements in this direction was obtained by considering, instead of a sphere, the smallest convex domain with rotational symmetry: this leaded to a more flexible approach to the problem [3].
So, let us consider the smallest convex domain Σ enclosing the scattering system, and let indicate with C' its intersection with a meridian plane Π (see fig. 5.3).

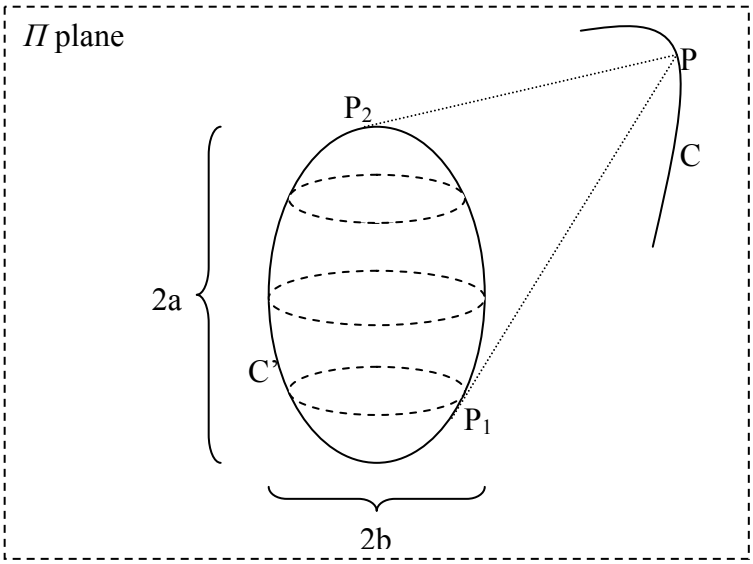


Figure 5.3 – Non-spherical scattering envelope

In [3], for and observation curve C laying on the plane Π, two expressions are found for the optimal field reduction and curve parameterization:

$$\gamma = \beta\left[\frac{R_1 + R_2}{2} + \frac{s'_1 - s'_2}{2}\right] + const \tag{5.20}$$

$$\xi = \frac{\beta}{W}\left[\frac{R_1 - R_2}{2} + \frac{s'_1 + s'_2}{2}\right] + const \tag{5.21}$$

where R_1 and R_2 are the distances between the observation point P and the two tangency points P_1 and P_2 on the source envelope (see fig. 5.3), s'_1 and s'_2 are the arclength coordinates of P_1 and P_2, respectively. It is important to note that the expressions (5.20) and (5.21) only depend on the observation point, and provide an (orthogonal) coordinate system on the plane Π. Moreover, when the

observation point P goes to infinity, γ tends to coincide with βr, i.e. a normalized radial coordinate, and ξ becomes an angle-like coordinate. In the case Σ is a spheroid of rotation, and Π a meridian plane, the coordinate curves of the system (γ,ξ) are shown to be the families of ellipses and hyperbolas confocal to C' (see figure 5.4) .

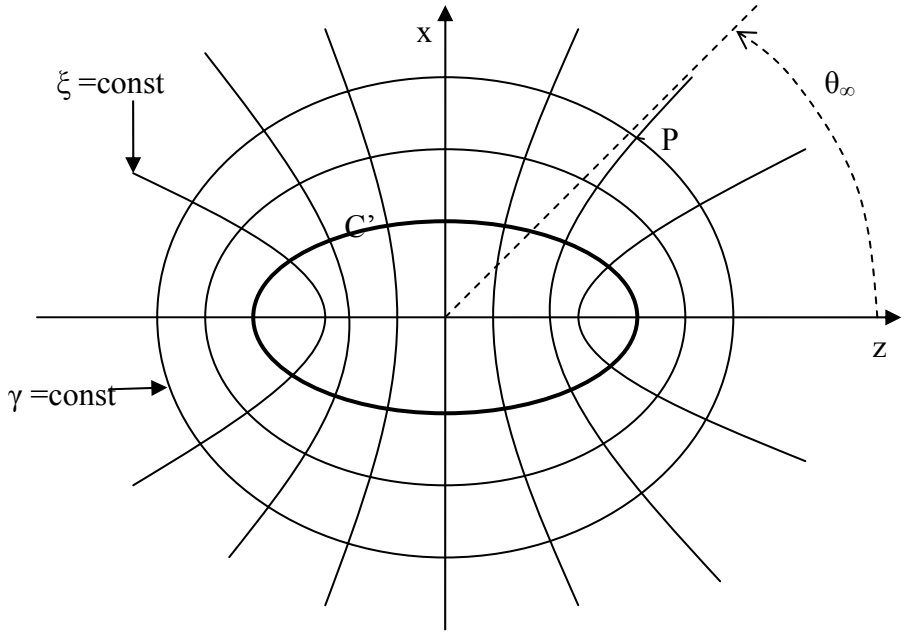


Figure 5.4 – Coordinate curves in the case of spheroidal scattering envelope

5.1.6 The degrees of freedom of the field

Important considerations on the degrees of freedom of the scattered fields can be derived from the application of the optimal sampling theory. When the observation point encircles the scattering system along an ellipse of length ℓ', ξ covers a 2π range, and the variation of the expression (5.21) is $\beta\ell'/W$, so that it can be chosen:

$$W = \beta\ell' / 2\pi \tag{5.22}$$

By performing a sampling at Nyquist rate $\Delta\xi = \pi / W$, the number of point to consider on any meridian curve enclosing the scattering system is

$$N_\xi = \beta\ell' W / \pi W = 2\ell' / \lambda \tag{5.23}$$

that can be considered as the number of the field's degrees of freedom on the plane Π. As seen previously, the band limitation error decays exponentially with the excess bandwidth, so that the field on the curve C can be approximated with a number of samples slightly larger than N_ξ, whatever is the size of the observation curve. Since the optimal sampling points on any observation curve are the intersections with the ξ=*constant* coordinate curves, we can say that, in the elliptic case, the N_ξ degrees of freedom of the field "moves along the hyperbolas" of the (γ,ξ) reference system.

The reasoning presented on the degrees of freedom on a closed curve on a meridian plane can be extended to any closed surface with rotational symmetry. In fact, the rotation of any point of the plane Π around the z-axis describes an azimuthal circumference, say C_φ. The field spatial bandwidth on C_φ can be estimated with analogous geometrical consideration [3], so that the number of samples needed to describe the field on a closed surface enclosing the scattering system turns out:

$$N \cong \frac{area\ of\ (\Sigma)}{(\lambda / 2)^2} \tag{5.24}$$

Since the tangential components of the radiated field on any closed surface including all the scatterers determine the field in any external point, we can say that the equation (5.24) provides the number of degrees of freedom of the field in the whole space.

5.2 Application to the A.S. positioning issue

Let us consider a PEC profile S, parameterized with the arc length s, illuminated by an incident field. As seen in chapter 3, the MAS technique is based on the choice of a set of auxiliary sources whose radiated field matches at best the boundary conditions over the scatterer profile: the fulfilment of this condition with a given precision ensures that the field radiated by the auxiliary sources outside the boundary agrees with the true scattered field within the same precision.

For large scatterers, according to the theory on the spatial bandwidth of the field, the field over S due to an arbitrary set of external primary sources (some wavelength far from it), as a function of the arc length, is essentially band limited to the free space propagation constant β. In the same way, if the distance of the auxiliary sources from the scatterer surface is small compared to its size, but not less than a few wavelengths, their field on S is also band limited to β. This ensures that if the incident field and that due to the ASs satisfy the boundary condition at a number N of uniformly spaced sampling points larger then the Nyquist number $N_0=L/(\lambda/2)$ (L being the length of S), this condition is satisfied all over S within the band limitation error, which decreases very rapidly with the oversampling ratio $\chi=N/N_0$. Thus, the field radiated by the auxiliary sources, within this error, coincides with the scattered one. As N_0 is essentially equals the number of degrees of freedom of the scattered field, the MAS should be able to achieve non redundancy and require a minimum possible number of auxiliary sources.

However, the location of the ASs is by no means a trivial task. In fact, they should lie well inside S, to make the band limitation error negligible, but not too much, to avoid a prohibitive ill-conditioning of the impedance matrix, which is also affected by the way they are distributed.

An answer to this question can be provided by the theory of optimal sampling representation of scattered fields, which involves a “natural” orthogonal coordinate system, strictly tied to the scatterer’s geometry, as shown in the previous paragraph. From this representation an ansatz follows: while the testing points are displaced uniformly at Nyquist rate on the scatterer profile, the ASs can be positioned inside

S, in correspondence with each test point, along the *ξ=constant* coordinate curves.

5.2.1 The case of an elliptic profile

Let us refer to the case of a 2D profile of elliptic shape. The natural system is the elliptic coordinate system confocal with the given profile. We propose to distribute the auxiliary sources on an interior coordinate curve (a confocal ellipse, in this case), in correspondence of the orthogonal coordinate curves (here confocal hyperboles) trough the sampling points (see Fig. 5.5) [20].

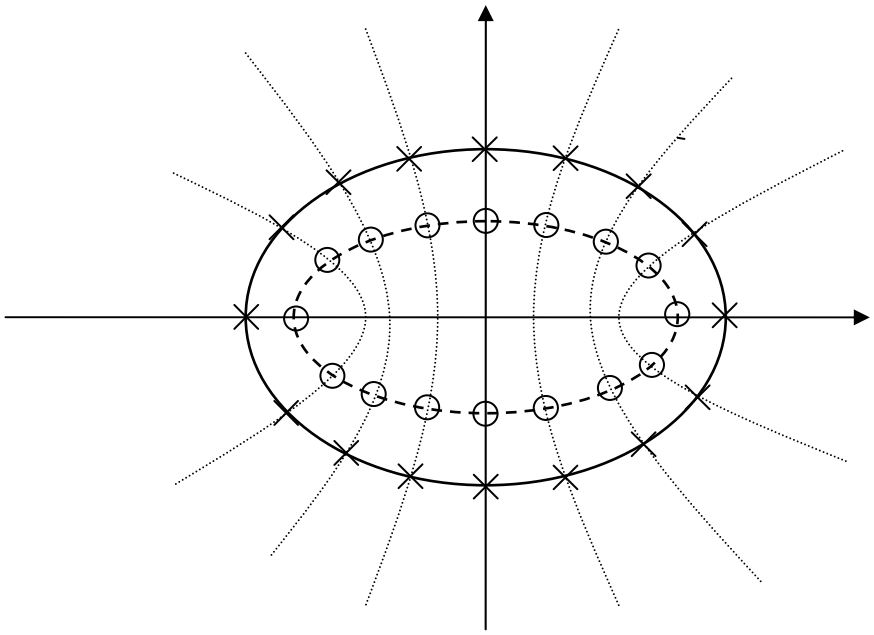


Figure 5.5 – Positioning of the auxiliary sources inside an elliptic scatterer. The scatterer profile is indicated with a solid line, the auxiliary surface with a dashed line, the confocal hyperboles with dotted lines, the testing points with crosses, the auxiliary sources with circles

This criterion can be applied in a similar way to the case of a different scattering geometry, by exploiting eqs. (5.20) and (5.21).

5.3 Improvements to the optimal positioning

In the MAS technique considered above, the testing points are uniformly distributed along the curvilinear abscissa on the scatterer surface, according to a spatial sampling of the incident field on S at a constant spatial frequency β. However, exploiting the theory on the sampling of the scattered fields, a further improvement can be obtained, and used to optimize the positioning of the MAS testing points.

In theory, in order to minimize the representation error we have to maximize the value of the parameter σ_m, which is a functional of the curve parameterization, appearing in the expression of the field sampling error given in §5.1. The problem of finding the curve parameterization (and, as a consequence, the optimal positions of the testing points) that maximizes the minimum of the σ parameter, is mathematically quite hard, and the solution can be found only for given geometries.

Anyway, in order to find a generic expression of a curve parameterization able to reduce the sampling error, an approach suggested by the professor O.M. Bucci is described and applied in the following.

Let us parameterize the boundary domain with the arc length s. Any given point $\underline{r}(s)$ on the curve is critical for a possible source position $\underline{r}'$ (in particular, it happens when the distance vector $\underline{r}$-$\underline{r}'$ is tangent to the boundary in $\underline{r}$), since w_0 reaches its maximum value β. In this condition, the σ parameter determining the error decay is written:

$$\sigma(\underline{r}') = \left(\frac{2}{|\dddot{\psi}|^{1/3}} (w - \beta) \right) \tag{5.25}$$

where $\left|\dddot{\psi}\right| = \left|\frac{\partial^3}{\partial \xi^3}\psi(\xi, \underline{r}')\right|_{\xi=\xi_0}$.

From the expression (5.25) it turns out that for a fixed value of the sampling bandwidth *w*, the σ parameter (and, as a consequence, the sampling error) is not constant, due to the variation of $\left|\dddot{\psi}\right|$. This means that the uniform distribution of the sampling points along the curvilinear abscissa provides a variable representation error. By an ansatz, aiming to a more convenient displacement of the testing points on the boundary, we can try to use a variable sampling bandwidth *w* to compensate the variation of the representation error, making the expression (5.25) constant. From a practical point of view, this will result in thickening of the testing points where the actual spatial bandwidth of the incident field is expected to be higher.

Concerning the evaluation of $\left|\dddot{\psi}\right|$, assuming constant the reduction function, we have:

$$\max_{\underline{r}'}\left|\frac{\partial^3}{\partial \xi^3}\psi(\xi, \underline{r}')\right|_{\xi=\xi_0} = \beta \max_{\underline{r}'}\left|\frac{\partial^3}{\partial \xi^3}R(\xi, \underline{r}')\right|_{\xi=\xi_0} \qquad (5.26)$$

Geometrical considerations allow to turn the above expression into:

$$\max_{\underline{r}'}\left|\frac{\partial^3}{\partial \xi^3}\psi(\xi, \underline{r}')\right|_{\xi=\xi_0} = \beta / \rho^2 \qquad (5.27)$$

where ρ is the local curvature radius of the curve. From (5.25) we obtain:

$$\sigma = \left(\frac{2\rho^2}{\beta}\right)^{1/3}(w - \beta) \qquad (5.28)$$

The relation shows that the σ parameter, with a fixed w, increases with the curvature of the observation domain. Aiming to make constant the error expression (5.13), we set:

$$\sigma^{3/2} = const$$

i.e., from (5.25)

$$\sigma^{3/2} = 2\beta\rho\left(\frac{w}{\beta}-1\right)^{3/2} = const \qquad (5.29)$$

Now, w/β is the ratio between two spatial frequencies and can be rewritten in terms of sampling rates as:

$$\chi = \frac{w}{\beta} = \frac{\lambda/2}{\Delta n/\Delta s} \qquad (5.30)$$

where Δs is the increase curvilinear abscissa and Δn the corresponding increase of the number of sampling points. Assuming $\Delta n/\Delta s \approx dn/ds$, we have from the eq. 5.29:

$$\frac{dn}{ds} = \frac{2}{\lambda} + \left(\frac{Y}{\rho}\right)^{2/3} \qquad (5.31)$$

The constant Y can be determined by integrating left and right hand over the whole observation curve (whose length is indicated with L):

$$\begin{aligned} N &= \int_0^N dn = \int_0^{L/\lambda}\left(2+\left(\frac{Y}{\rho}\right)^{2/3}\right)d\xi = \\ &= N_0 + Y^{2/3}\int_0^{L/\lambda}\frac{1}{\rho^{2/3}}d\xi \end{aligned} \qquad (5.32)$$

where $\xi = s/\lambda$ and $N_0 = 2L/\lambda$, i.e.,

$$\frac{N-N_0}{N_0}=\frac{Y^{2/3}}{2L/\lambda}\int_0^{L/\lambda}\frac{1}{\rho^{2/3}}d\xi=\frac{Y^{2/3}}{2}\left\langle k^{2/3}\right\rangle \qquad (5.33)$$

$k=1/\rho$ being the local curvature, and $\left\langle k^{2/3}\right\rangle=\frac{\lambda}{L}\int_0^{L/\lambda}k^{2/3}d\xi$ representing the average value of $k^{2/3}$ along the curve. Since $N/N_0=\chi$, we have:

$$Y=\left(2(\chi-1)\frac{1}{\left\langle k^{2/3}\right\rangle}\right) \qquad (5.34)$$

Substituting the value of Y into the equation (5.31) we find:

$$\frac{dn}{d\xi}=2\left(1+(\chi-1)\frac{k^{2/3}}{\left\langle k^{2/3}\right\rangle}\right) \qquad (5.35)$$

This equation (5.35) suggests, for a given value of the ratio χ, the optimal positioning of the $N=2L\chi/\lambda$ sampling points on the scatterer's surface. In fact it is easy to find the following parameterization for the curve:

$$\begin{aligned}\frac{n}{N}&=\frac{1}{N}\int_0^n dn=\frac{\lambda}{\chi L}\int_0^{s/\lambda}1+(\chi-1)\frac{k^{2/3}}{\left\langle k^{2/3}\right\rangle}d\xi=\\&=\frac{s}{\chi L}+\frac{\lambda}{L}\left(1-\frac{1}{\chi}\right)\frac{1}{\left\langle k^{2/3}\right\rangle}\int_0^{\xi}k^{2/3}d\xi\end{aligned} \qquad (5.36)$$

The value of n given by expression can be evaluated numerically or analytically, depending on the characteristic of the domain curve.

5.3.1 Application to the elliptic geometry

The results of the previous Section can be applied to the positioning of the MAS testing points according to the expression (5.36), by taking into account the bandwidth increase due to the profile curvature. In particular, a more effective positioning criterion for the testing point and, as a consequence, a reduced redundancy of the auxiliary sources technique is expected.

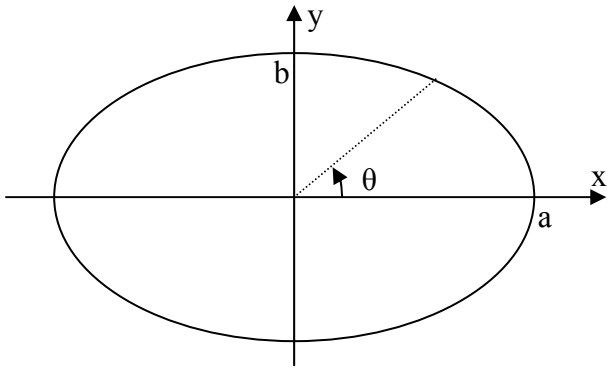


Figure 5.6 – Elliptic scatterer

Let us consider an ellipse of semiaxes (*a,b*) (see figure 5.6). The curvature of this profile is found to be:

$$k = \frac{ab}{\left(a^2 \sin^2(\theta) + b^2 \cos^2(\theta)\right)^{3/2}} \tag{5.37}$$

For the average of $k^{2/3}$ can be written:

$$\begin{aligned} \left\langle k^{3/2} \right\rangle &= \frac{\lambda}{L} \int_0^{L/\lambda} k^{3/2} d\xi = \\ &= \frac{1}{L} \int_0^{2\pi} \frac{(ab)^{2/3} d\theta}{b\sqrt{1-(1-e^2)\sin^2(\theta)}} = \frac{(ab)^{2/3}}{Lb} 4K(m) \end{aligned} \tag{5.38}$$

$K(e)$ being the complete elliptic integral of the first kind, with parameter $e^2 = 1-(a^2/b^2)$. Analogously, we find:

$$\int_0^{\xi} k^{2/3} d\xi = \int_0^{\theta} k^{2/3} \frac{d\xi}{d\theta} d\theta = \\ = \int_0^{\theta} \frac{(ab)^{2/3} d\theta}{\lambda b \sqrt{1-(1-e^2)\sin^2(\theta)}} = \frac{(ab)^{2/3}}{\lambda b} F(\theta, m) \tag{5.39}$$

according to the relation $\frac{d\xi}{d\theta} = \frac{1}{\lambda}\sqrt{a^2 \sin^2(\theta) + b^2 \cos^2(\theta)}$, where $F(\theta,e)$ is the incomplete elliptic integral of the first kind.
By substituting (3.38) and (5.39) in (5.36), the optimal parameterization of the elliptic profile is found:

$$\frac{n}{N} = \frac{s}{\chi L} + \left(1 - \frac{1}{\chi}\right) \frac{F(\theta, e)}{4K(e)} \tag{5.40}$$

As a conclusion, in the 2D elliptic scatterer case, the optimal sampling of the profile can be performed by taking into account the actual variation of the local bandwidth of the field due to the curvature and exploiting the analytical expression (5.40). In the following chapter, the main results of a wide numerical analysis will confirm the validity of this criterion.

Chapter 6

Numerical Analysis

The main results of the numerical analysis performed to enlighten the performance of the technique are presented in this Chapter. The aim is to show the improvement of the performances of the Method of Auxiliary Sources, with respect to the non-redundancy of the number of unknowns and the required computational effort.

6.1 Scattering by a circular cylinder

First, the scattering by a PEC indefinite cylinder of circular section is considered. This simple geometry allows to focus the analysis on some relevant parameters, and to make comparison with the available analytical solution of the scattering problem, thus giving an accurate evaluation of the performances of the approach.
In this case the determination of the positioning of the auxiliary source (AS) and collocation point (CP) is almost trivial. In fact, the radial symmetry of the geometry imply that the spatial bandwidth of the field along the boundary is constant and, as a consequence, the CPs are displaced uniformly with respect to the curvilinear abscissa. Analogously, the auxiliary surface is a concentric cylinder, and the ASs are placed uniformly in correspondence with the testing points. The only parameters to be set are:

1) the "dept" of the ASs beneath the scattering surface, i.e. the radius of the auxiliary cylinder.
2) the "oversampling" factor χ involved in the spatial frequency setting for the AS and CP positioning.

Let us denote with a and a_0 the radius of the PEC and auxiliary cylinder, respectively, and with χ the "oversampling" factor:

$$\chi = \frac{N\lambda}{4\pi a} \tag{6.1}$$

N being the number of the ASs (Fig 6.1).
As seen before, the problem consists in evaluating the Auxiliary sources excitations whose radiated field matches the boundary condition, i.e., satisfy the equation:

$$\underline{\underline{Z}} * \underline{I} = -\underline{E}_{inc} \tag{6.2}$$

where $\underline{\underline{Z}}$ is the impedance matrix, $\underline{I}$ the AS excitation vector and $\underline{E}_{\text{inc}}$ the vector containing the value of the tangential components of the incident field on the CPs.
The field produced by the *n*_th AS on the *m*_th CP is represented by the matrix element:

$$z_{m,n} = -\frac{j}{4} H_0^{(2)}(k \mid r_m - r'_n \mid) \tag{6.3}$$

where $H_0^{(2)}$ is the Hankel function of second kind and order zero and $\underline{r}_{\text{m}}$ and $\underline{r}_{\text{n}}$ the position vectors of the *m*_th CP and *n*_th AS points, respectively. By applying the addition theorem [61], eq. (6.3) can be rewritten as:

$$z_{m,n} = -\frac{j}{4} \sum_{l=-\infty}^{\infty} J_l(ka_0) H_l^{(2)}(ka) \exp(-jl(\phi_m - \phi'_n)) \tag{6.4}$$

where ϕ_m and ϕ'_n are the angles of the *m*_th CP and of the *n*_th AS, respectively. Exploiting the expression (6.4) is possible to diagonalize and analytically solve the linear system (6.2) [47]. The eigenvalues of the impedance matrix are:

$$\lambda_q = -\frac{jN}{4}\sum_{s=-\infty}^{\infty} J_{q+sN}(ka_0)H^{(2)}_{q+sN}(ka) \tag{6.5}$$

and the eigenvectors:

$$\underline{g}_q = -\frac{1}{\sqrt{N}}\left[\exp(-jq\phi_1),...,\exp(-jq\phi_N)\right] \tag{6.6}$$

q assuming values between 1 and N.
The condition number of the impedance matrix strongly influences the stability and the accuracy of the numerical solution. In this case, it depends on the radii a ed a_0, and on the number of CPs-Ass, according to the relationship:

$$\kappa(a_0, a, N) = \frac{\max_q\left|\lambda_q\right|}{\min_q\left|\lambda_q\right|} = \frac{\max_q\left|\sum_{s=-\infty}^{\infty} J_{q+sN}(ka_0)H^{(2)}_{q+sN}(ka)\right|}{\min_q\left|\sum_{s=-\infty}^{\infty} J_{q+sN}(ka_0)H^{(2)}_{q+sN}(ka)\right|} \tag{6.7}$$

By means of an asymptotical analysis [47], for large scatterers, the behaviour of the condition number for increasing N can be estimated as:

$$\kappa \propto \frac{\lambda N}{2\pi a}\left(\frac{a_0}{a}\right)^{-N/2} \tag{6.8}$$

Eq. (6.8) shows that, due to the relationship $a_0<a$, the condition number grows exponentially with the number of auxiliary sources N. This is an issue typical of the MAS approach not encountered in the classical MoM implementations wherein the equivalent sources are placed on the boundary surface, and the impedance matrix does not suffer of so severe ill-conditioning.

Another point of interest concerns the presence of some "resonance phenomena" already observed in some MAS implementations [47]. In fact, for high values of N ($N>>ka_0$) the infinite sums in (6.7) can be approximated by means of the only $s=0$ term, leading to the approximate expression for the condition number:

$$\kappa(a_0, a, N) \approx \frac{\max_q \left| J_q(ka_0) H_q^{(2)}(ka) \right|}{\min_q \left| J_q(ka_0) H_q^{(2)}(ka) \right|} \tag{6.9}$$

It is noted that when the value of ka_0 approaches a zero of the Bessel functions, even if the $\frac{a_0}{a}$ ratio is close to 1, the condition number steadily increases and the numerical solution exhibits an outstanding error.

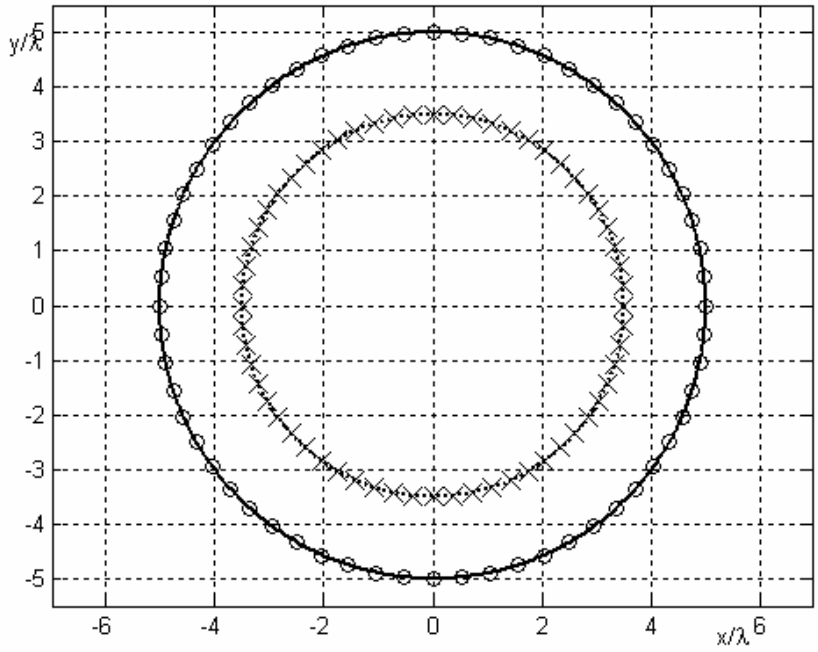


Figure 6.1 – Circular cylinder in 2D. The continuous line shows the boundary of the scatterer, the dotted line indicates the auxiliary surface. Circle and crosses shows the positions of the collocation points and of the auxiliary sources, respectively.

According to equation (6.8), to obtain a well conditioned impedance matrix, the ratio a_0/a should approach the value 1. However, the fact that both the incident and the AS fields, in the non redundant implementation of the MAS previously presented, should be essentially band-limited to β does not allows values of (a_0/a) too close to unity. Accordingly, differences between a and a_0 at least equal to 1-2λ should be ensured. Provided that this distance is large enough, the error of the proposed MAS technique will show the exponential behaviour discussed in chapter 5, i.e., in the case of a circular cylinder of radius a, the error should be asymptotically proportional to:

$$\exp(-C\beta a(N/N_0 - 1)^{3/2}) \tag{6.10}$$

being $N_0=2\beta a$, and C a constant depending on the geometry. The ratio N/N_0 is also referred as the "oversampling ratio" and is indicated with χ.

Two cases have been numerically worked out (both with a=5λ) by considering a plane wave and a point-wise source. The numerical analysis has been carried out by varying the "oversampling" ratio χ and the "penetration" ratio $v=a_0/a$ and evaluating the errors of the solutions as compared to the known analytical solution.

Let us consider a Cartesian reference system with the z-axis equal to the cylinder axis and an incident plane wave propagating along the x-axis, i.e., $E_{inc}=i_z exp(-jkx)$. The analytical solution is given by [68]:

$$E_z^s(\rho,\varphi) = -\sum_{n=0}^{\infty} \varepsilon_n (-i)^n \frac{J_n(ka)}{H_n^{(2)}(ka)} H_n^{(2)}(k\rho)\cos(n\phi) \tag{6.11}$$

ε_n being the Neumann symbol ($\varepsilon_0=1$; $\varepsilon_{n\neq 0}=2$), $H_n^{(2)}$ being the Hankel function of second kind and order n, J_n being the Bessel function of order n, and (ρ,φ) being the polar coordinates of the observation point with respect to the centre of the cylinder.

For the sake of numerical evaluation, the infinite sum in (6.11) can be truncated to an index $\tilde{n} > k\rho$, according to the fact that the values of the Bessel function become negligible when the order exceeds the argument.

The modulus of the MAS impedance matrix obtained by setting $\chi=1.5$ and $(a_0/a)=0.7$ is shown in figure 6.2. As can be seen, it is a dense one, even if the off-diagonal elements have smaller modulus with respect to the diagonal ones. Furthermore, due to the symmetry of the geometry, the N-by-N matrix has a *circulant structure*, i.e., its rows are composed by cyclically shifted versions of a sequence of N values.

The eigenvalues of the matrix obtained by a Singular Value Decomposition, are shown in fig. 6.3 as a function of the radius of the

auxiliary surface. As expected, they decay exponentially as the index *n* exceeds the number of degrees of freedom of the scattered field, i.e., $N_0=4\pi a_0/\lambda$. Moreover, the condition number increases according to the law (6.8) as a function of the ratio (a_0/a).

The normalized main square difference between the exact scattered field and the one estimated by means of the MAS, in the case of an incident plane wave, is reported under Fig.6.4 as a function of χ and a_0/a. As expected, the error behaviour approaches the asymptotical behaviour (6.10) as the (a_0/a) ratio decreases, due to the fact that the distance between the sources and the collocation points becomes large enough to assume that the radiated field is a band-limited function. In this "saturation" region, the error due to the MAS approach has the same level of the minimal band limitation error associated to the representation of the field by means of *N* samples, according with the optimal sampling theory [1]. At the same time, as the ASs shrink, the ill conditioning of the impedance matrix increases, so that, for small values of (a_0/a), the error curve may explode for relatively low values of χ before attaining a satisfactory error level. The error explosion occurs when the ill-conditioning becomes too high with respect to the numerical accuracy of the used computational tool. As this happens, the numerical error introduced by the finite machine precision is amplified and start to prevail on the error due to the MAS formulation. The results do no show any "resonance phenomena" found by some researchers and explained with the previous considerations related to equation (6.9). It is believed that this is due to the fact that the number *N* of used unknowns is not so much larger than ka_0, essentially equal to the number of degrees of freedom of the problem, so that, in this case, the approximations leading to eq. (6.9) are not justified.

Another interesting result concerns the dependence of the error on the dimensions of the scatterer . According to the asymptotical law (6.10), the error decreases more rapidly as the size of the cylinder increase. This behaviour is confirmed by the numerical results reported in Figure 6.5, wherein the analytically forecasted error and the actual one, obtained by a MAS implementation working in the saturation area, is reported in the cases $a=5\lambda$, $a=10\lambda$, $a=20\lambda$ and $a=30\lambda$.

According to these considerations, a rule for positioning of the ASs inside the cylinder should take into account two contrasting demands:

the distance (a-a_0) should be large enough to work in the "saturation" region and small enough to control the ill-conditioning, increasing when the auxiliary sources shrink. It is important to note that, once the "saturation" has been reached, that is to say the minimum attainable error has been obtained, any further increase of the distance (a -a_0), only will raise the condition number of the matrix.

The parameters of a MAS technique can be set by referring to the abacus of fig. 6.5, giving the value of χ required to ensure the prescribed error level for the assigned problem size. Successively, the distance (a -a_0) must be chosen large enough to ensure that the field radiated by the auxiliary sources on the scatterers boundary is essentially a band-limited function. Finally, eq. (6.8) allows to estimate the condition number of the impedance matrix, so that its compatibility with the numerical precision of the used tools can be estimated.

The numerical analysis has shown that, for large scatterers, using standard computers featuring the "double precision" arithmetic, results with a relative error of -100dB can be achieved before the effect of the ill-conditioning becomes relevant.

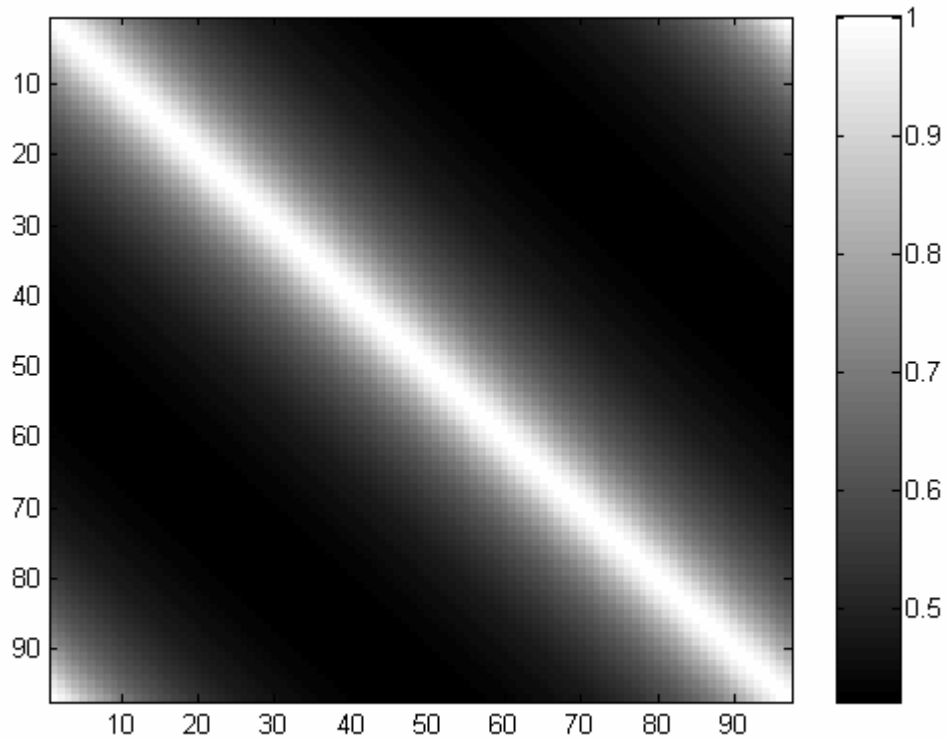


Figure 6.2 – MAS impedance matrix, in modulus, for a circular cylinder of radius 5λ, for χ=1.5 and (a_0/a)=0.7.

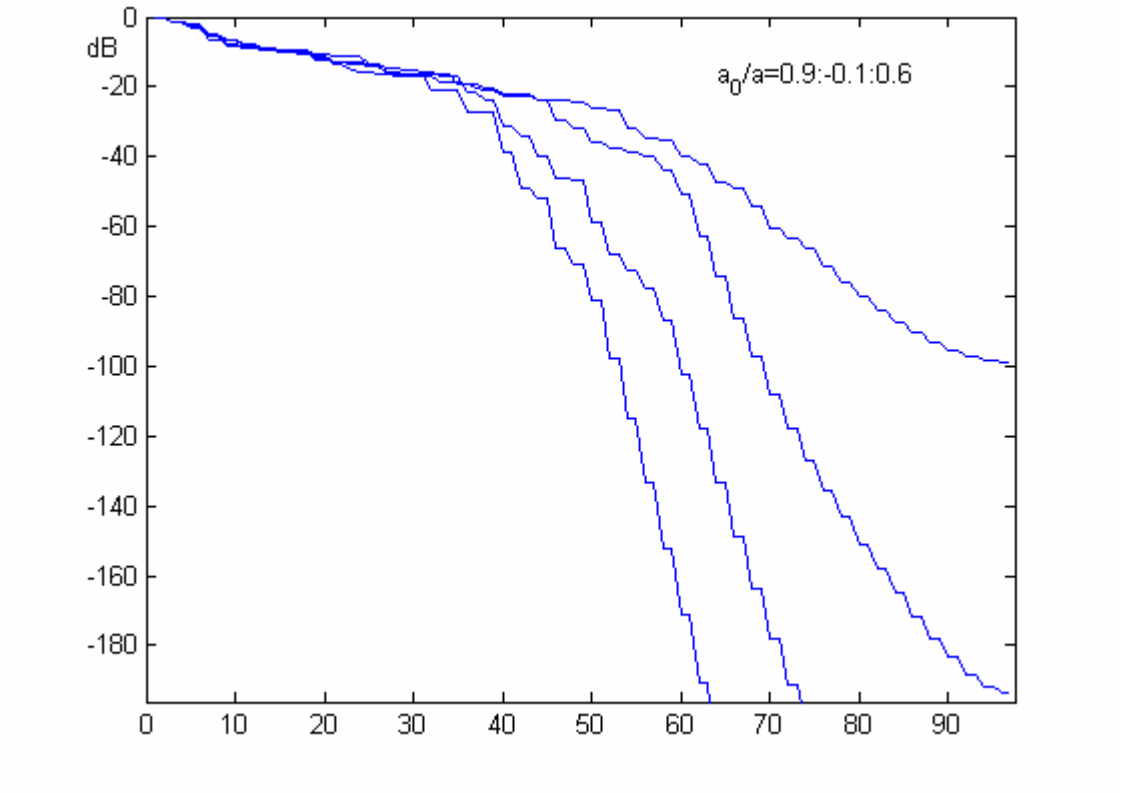


Figure 6.3 – Behaviour of the sequence of the normalized singular values for a circular cylinder of radius a, for various radii of the auxiliary surface.

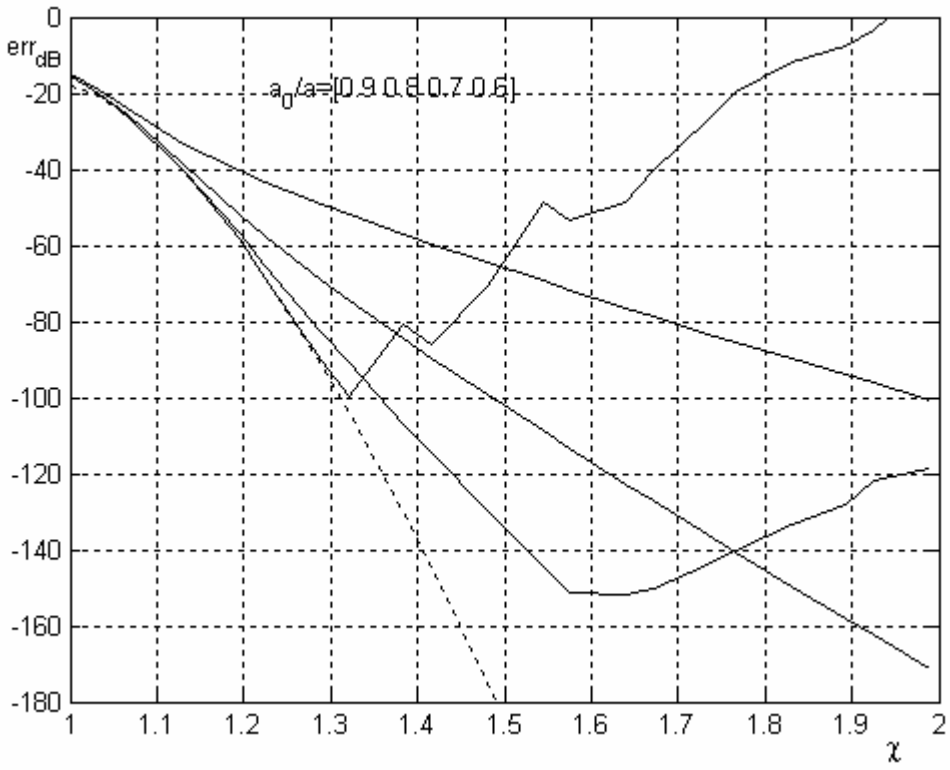


Figure 6.4 – Relative error in scattered field evaluated via the MAS, for a cylinder of radius $a=5\lambda$ under plane wave incidence.

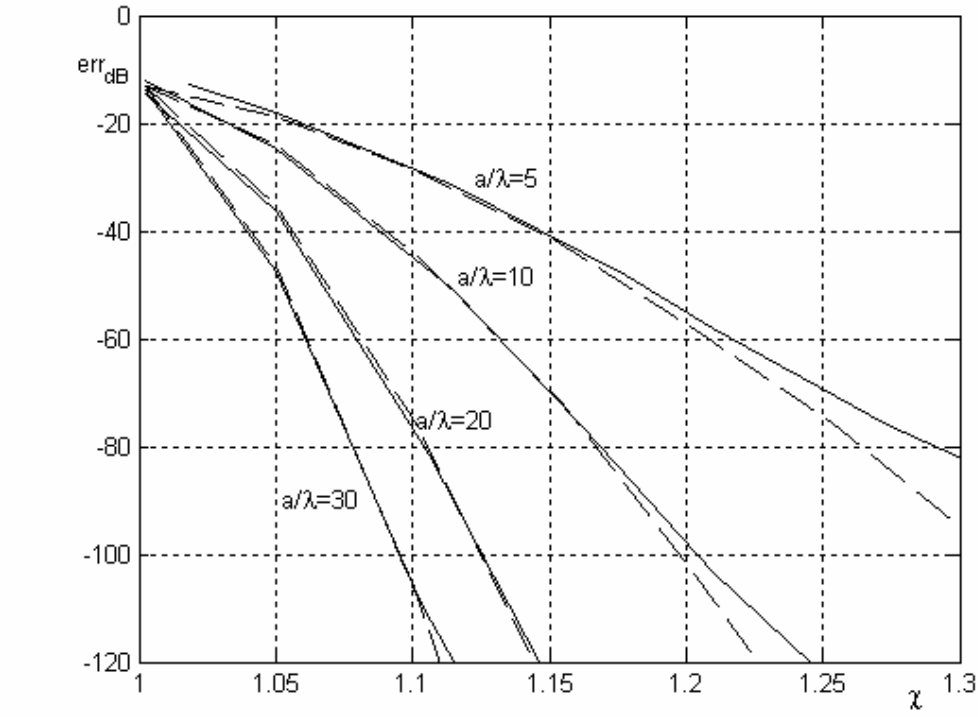


Figure 6.5 – Abacus reporting the relative error for various values of the radius a (solid lines), compared with the asymptotical behaviour of the band-limitation error (dashed lines)

6.2. Scattering by elliptical cylinders.

An elliptical scattering profile, without the radial symmetry of the circular cylinder, allows to study the effectiveness of the optimal positioning strategies based on the concept of the spatial bandwidth of the scattered fields. Since the radial symmetry is lost, the impedance matrix has no more a circulant structure, and it is not possible to speed up the evaluation and the inversion of the matrix by trivially exploiting the symmetry of the geometry.

Moreover, at variance of the circular scatterer case wherein only the oversampling ratio and the "dept" of the auxiliary surface are of concern, in the elliptical scatterer case the positioning of the ASs is less trivial. In the chapter 5, it was shown that the spatial bandwidth of the scattered field on the scatterer surface is limited to β. Furthermore, it was noted that any scattering geometry induces a "natural" coordinate system acting as a "grid" for the optimal sampling of the field. In a pictorial sense, it can be said that the degrees of freedom "move" along well known coordinate curves. Accordingly, it is "natural" to locate the auxiliary sources at the intersection of an interior coordinate curve with the orthogonal coordinate curves passing trough the testing points (see fig. 5.5).
In the case of the elliptic profile, the coordinate curves are sets of confocal ellipses and hyperbolas.
As known, the field scattered by a PEC elliptic cylinder has been evaluated in [68]. The elliptic cylindrical coordinates (*u,v*) are related to the Cartesian ones (*x,y*) by the transformation (see Fig. 6.6):

$$\begin{aligned} x &= f \cosh u \cos v \\ y &= f \sinh u \sin v \end{aligned} \tag{6.12}$$

where *2f* is the inter-focal distance. Moreover, defining

$$\begin{aligned} \xi &= \cosh u \\ \eta &= \cos v \end{aligned} \tag{6.13}$$

with $1 \le \xi < \infty$ and $-1 \le \eta \le 1$ the equation *u=constant* (as well as ξ*=constant*) defines an ellipse of foci (+*f*,0) and (-*f*,0), and eccentricity ξ^{-1}.

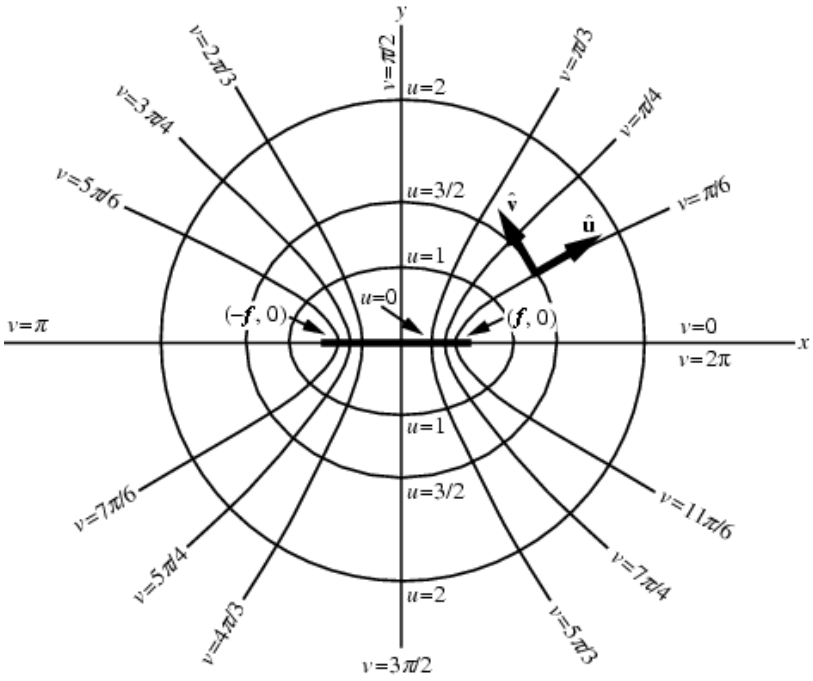


Figure 6.6 – Elliptic coordinate system

Let us consider a PEC scattering cylinder whose profile has equation $u=u_1$ and an incident plane wave propagating in the x-y plane at an angle ϕ_0 with respect to the negative x-axis. The incident field is given by:

$$E^i = \hat{z}\exp\left(-j\beta\left(x\cos\phi_0 + y\sin\phi_0\right)\right) \qquad (6.14)$$

and the scattered field is:

$$E_z^s = -\sqrt{8\pi}\sum_{m=0}^{\infty}(-j)^m \cdot$$
$$\cdot\left[\frac{1}{N_m^{(e)}}\frac{Re_m^{(1)}(c,\xi_1)}{Re_m^{(3)}(c,\xi_1)}Re_m^{(3)}(c,\xi)Se_m(c,\cos\phi_0)Se_m(c,\eta)+\right. \qquad (6.15)$$
$$\left.+\frac{1}{N_m^{(o)}}\frac{Ro_m^{(1)}(c,\xi_1)}{Ro_m^{(3)}(c,\xi_1)}Ro_m^{(3)}(c,\xi)So_m(c,\cos\phi_0)So_m(c,\eta)\right]$$

where $c = \beta f$, $\xi_1 = \cosh u_1$, $Re, o_m^{(1,3)}$ denotes the even and odd radial Mathieu function of the first and third kind, respectively, and Se, o_m is the angular even or odd angular Mathieu function [69]. The expression (6.15) can be evaluated by truncating the summation to a finite value, and the result can be used as a reference to estimate the accuracy of the MoM-like approaches.

6.2.1 A small elliptical cylinder, comparison between MAS and MoM

As a first case, we consider a PEC scatterer with major and minor axes equal to 10λ and 6λ, respectively, illuminated by an incident plane wave with a TM polarization propagating along the major axis. The small size of scatterer allow us to solve the scattering problem with a classical MoM techniques, and to compare its effectiveness with a MAS based technique.

The mean square error between the far field calculated with the MAS and the analytically known solution (6.15) is shown in Fig.6.8, as a function of the oversampling ratio χ, for different values of the ratio, ν, between the major axis of the auxiliary ellipse and the boundary length. As can be seen, when the distance between the auxiliary ellipse and the scatterer boundary increases, the error saturates. This is due to the fact that, when the sources are located far enough from the

scatterer's profile, the bandlimitation error of the corresponding radiated field becomes negligible as compared with that of the incident field. This is confirmed by the behaviour of the limit curve, which is proportional to $(N/N_0\text{-}1)^{3/2}$, matching the theoretical behaviour of the band limitation error previously determined.

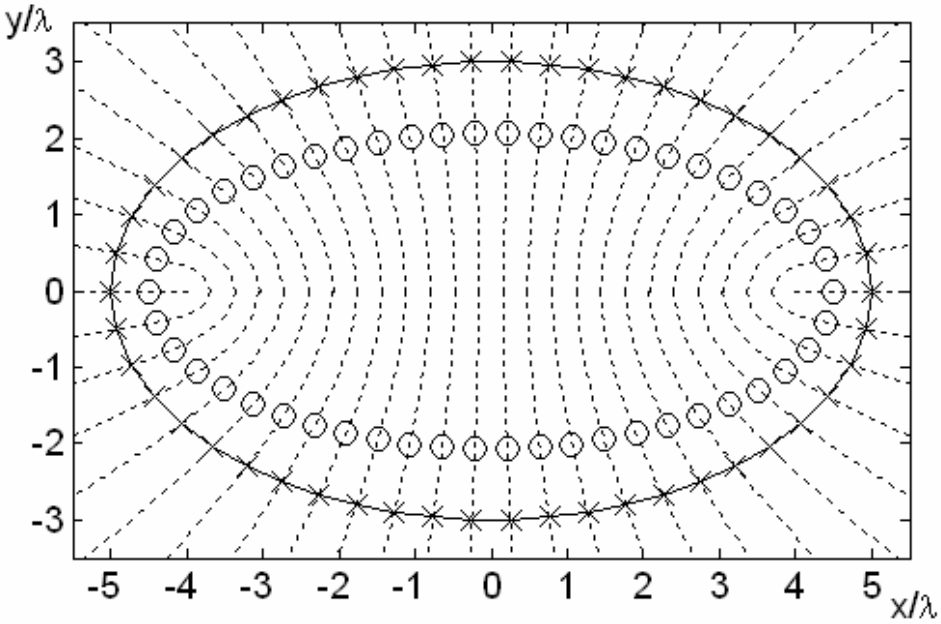


Figure 6.7 - Geometry of the problem. Solid line: scatterer profile; dotted lines: confocal hyperboles; crosses: testing points; circles: auxiliary sources.

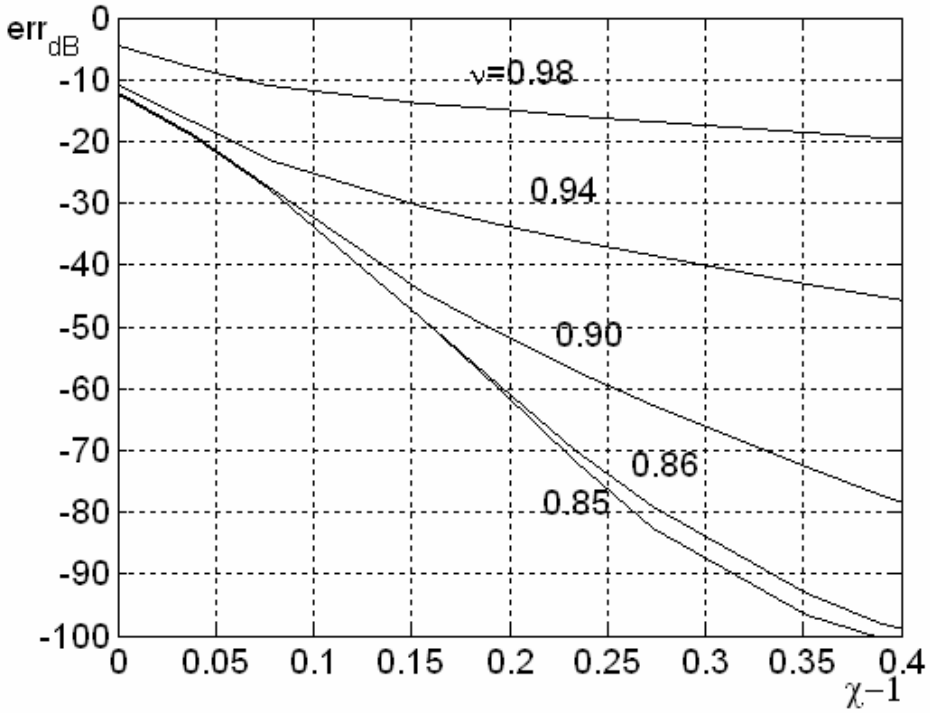


Figure 6.8 – Mean square error of the field evaluated by MAS, as a function of ν.

The figure 6.8 shows that, even for not too large scatterers, an oversampling ratio of the order of 20% ensures a negligible error, i.e., an error smaller than -60dB.
For the sake of comparison, a classical MoM based algorithm, exploiting triangular and pulse functions as basis and test functions, respectively, has been developed to solve the same scattering problem. At variance of MAS, filling the impedance matrix requires the numerical evaluation of integrals (instead of the direct computation of the interaction between 2 points).
Analogously to the MAS case, the numerically evaluated field has been compared with the exact one, and the error has been plotted with

respect to the parameter $\chi=N/N_0$, N being now the number of unknowns (basis and testing function) (Fig. 6.9).

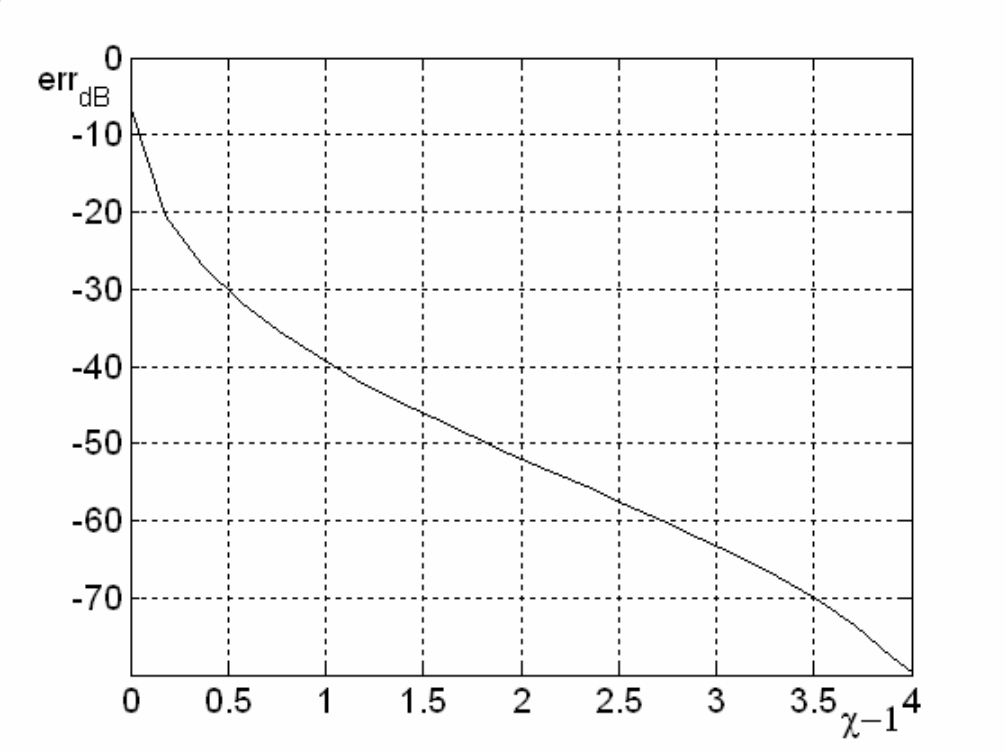


Figure 6.9 – Error of the scattered field evaluated with a standard MoM algorithm, for varying values of χ (ratio between the number of unknowns and the number of the field's degrees of freedom)

The figure 6.9 shows that, despite its higher complexity, the MoM algorithm needs a number of unknowns per wavelength much larger than our MAS technique in order to obtain the same error. For example, if we want to evaluate the scattered field with an error lower than -70 dB, we can choose ν=0.85 and χ=1.23 in our MAS algorithm (2.46 unknowns per wavelength), while we need at least χ=4.5 (9 unknowns per wavelength) to obtain the same precision with the

standard MoM. For large scatterers, this reduction in the number of unknows can be very significant, and allows the study of problems that otherwise would be prohibitive.

6.2.2 A set of two elliptical cylinders

Another case of interest concerns a scattering system made by more than one PEC object. We consider two "small" ellipses, with different eccentricity, illuminated by a plane wave and placed not so far by each other, so that the interaction between them cannot be neglected. Since no analytical result is available for such a problem, we are forced to exploit a classical MoM approach for comparison purposes.
In particular, we consider two ellipses with major and minor semi-axis equal to 5λ and 3λ and 3λ and 2λ, respectively (see Figure 6.10).

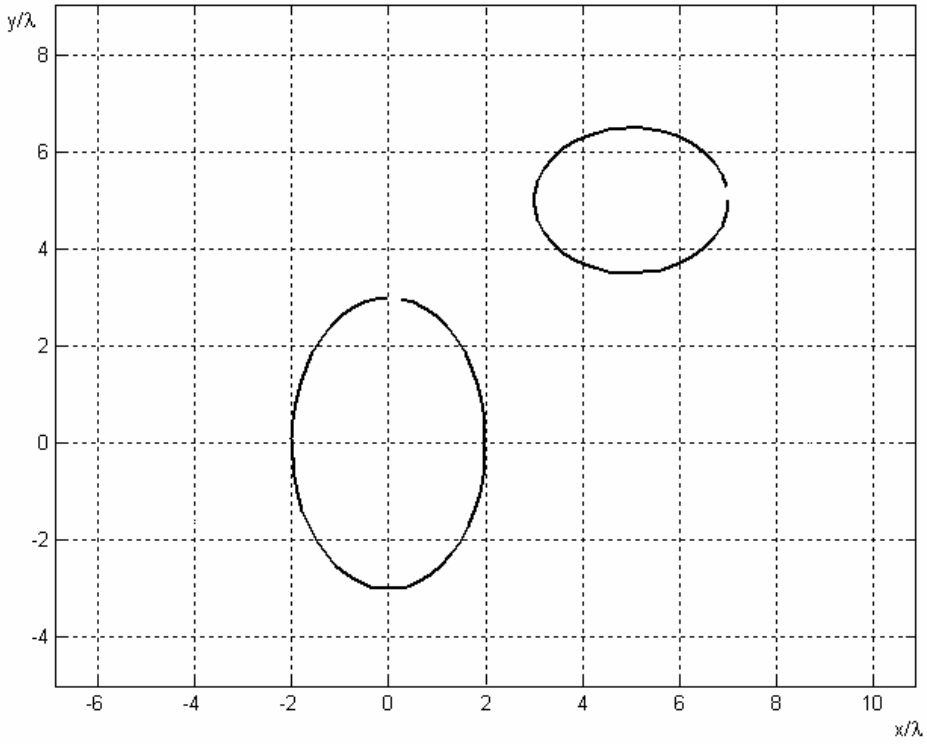


Figure 6.10 – Geometry of the scattering system

For each ellipse, the collocation points and the auxiliary sources have been located following the guidelines already established in the case of a single elliptic cylinder. In order to evaluate the accuracy of the solution obtained by applying the MAS based technique, the MoM algorithm with 20 unknowns per wavelength, and basis functions involving 4 interpolation points has been considered [70].
The mean square error is reported in Figure 6.11 as a function of the oversampling ratio χ, for various values of the ratio ν_0 between the major axes of the auxiliary ellipse and the length of the PEC boundary. Since the accuracy of the reference MoM field is not better than -47dB, the error can be analyzed just on a restricted range, and

the agreement with the expected exponential behaviour cannot be confirmed. Anyway, it is possible to confirm two aspects highlighted in the previous examples: the curves tend to saturate when the depth of the ASs reaches 1-2 λ inside the profile; the MAS error at saturation decays rapidly, and it is possible to obtain a good precision for values of oversampling slightly larger than the unity. An oversampling factor χ=1.2 ensures an error lower that -40dB.

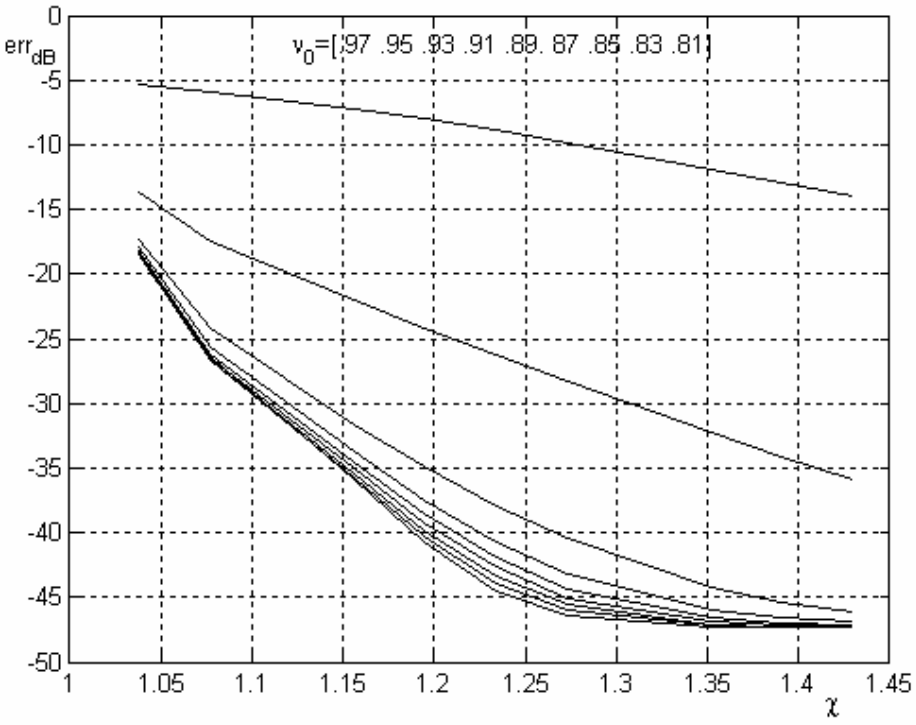


Figure 6.11 – Error of the MAS field, with respect to the MoM solution, for various values of the ratio between auxiliary and profile ellipse major axes.

6.2.3 Larger elliptical cylinders

The numerical analysis performed in the case of the circular cylinder (see Figure 6.5) shows that, for a fixed oversampling ratio, the relative error of the MAS based solution decreases with the size of the scatterer. As will be seen, the same behaviour holds also in the elliptic case wherein the angle of incidence ϕ_0 must be taken into account.

First, the numerical problems involved in the evaluation of the analytical expression of the scattered field (6.15) has been considered. In the case of large scatterers the number of terms needed to evaluate the summation in (6.15) seriously increase, and the attainment of the required accuracy became difficult.

As a matter of fact, to build a reference scattered field and to estimate the error of the MAS based solution, we use classical MoM algorithm with a strong increase of the number of the unknowns. In particular, to ensure an accuracy of -100dB up to 30 rooftop basis functions per wavelength have been used.

Three scattering elliptical cylinders of different size and three angles of incidence have been considered for a total of 9 cases. The values of the major and minor semi-axes are equal to 5λ, 3λ, 10λ, 6λ and 20λ, 12λ, respectively while the values of the incidence angle ϕ_0 are equal to 0, 45° and 90°, respectively. The resulting error of the field evaluated by the MAS with parameters guaranteeing the so called "saturation" condition, are shown in figures 6.12, 6.13 and 6.14.

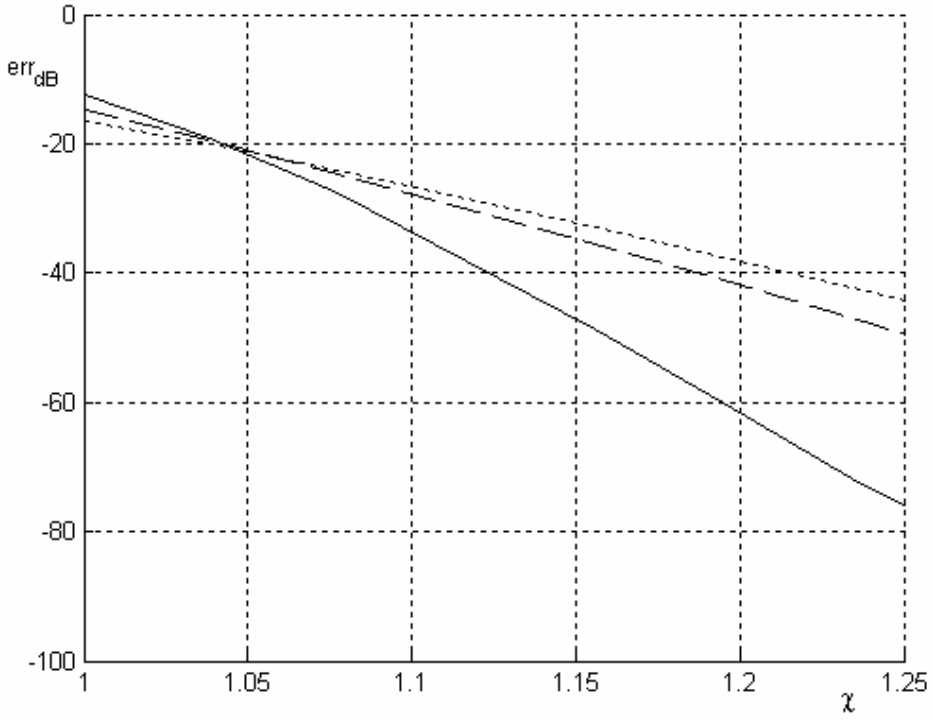


Figure 6.12 – Error of the scattered field evaluated by the MAS, with respect to the MoM solution, for a scattering elliptical cylinder of semiaxes (5λ, 3λ), under the following conditions of incident plane wave: ϕ_0 =0 (solid line), ϕ_0 =45° (dashed line), ϕ_0 =90° (dotted line)

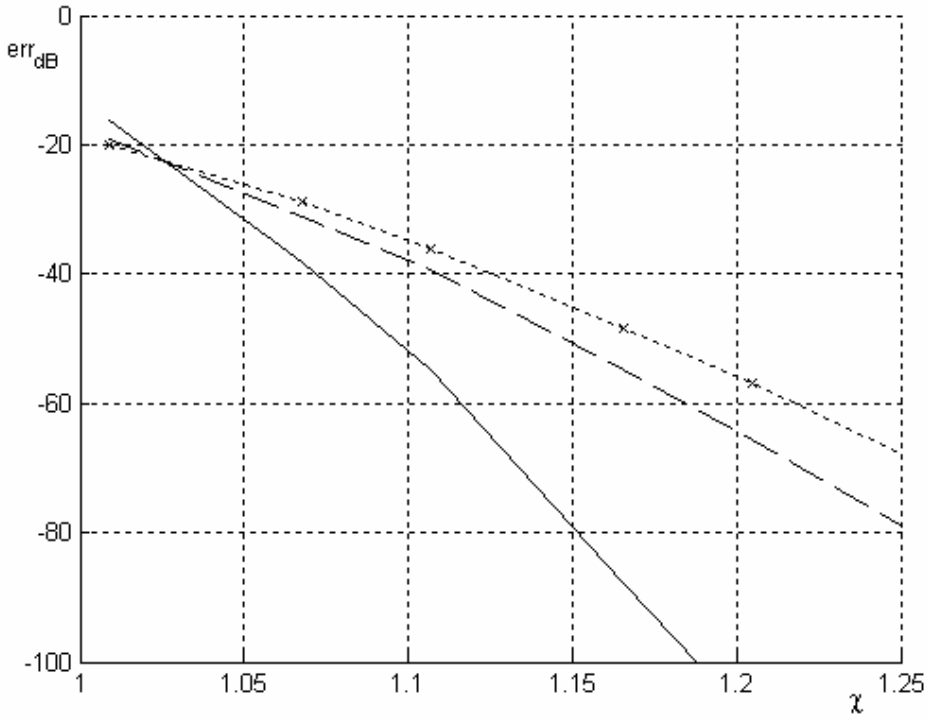


Figure 6.13 – Error of the scattered field evaluated by the MAS, with respect to the MoM solution, for a scattering elliptical cylinder of semiaxes (10λ, 6λ), under the following conditions of incident plane wave: ϕ_0 =0 (solid line), ϕ_0 =45° (dashed line), ϕ_0 =90° (dotted line)

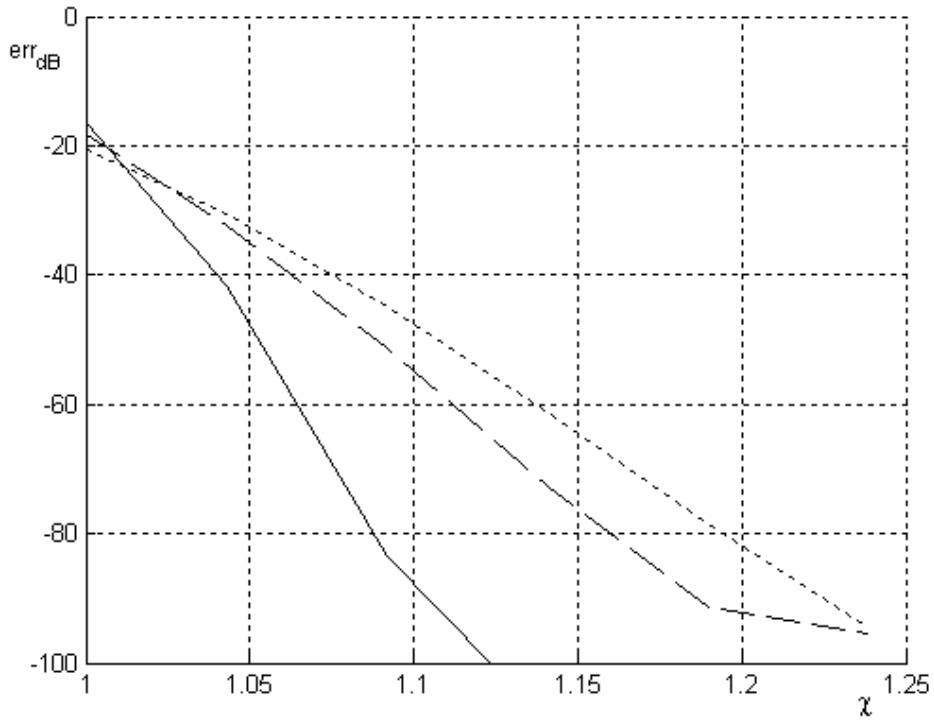


Figure 6.14 – Error of the scattered field evaluated by the MAS, with respect to the MoM solution, for a scattering elliptical cylinder of semiaxes (20λ, 12λ), under the following conditions of incident plane wave: ϕ_0 =0 (solid line), ϕ_0 =45° (dashed line), ϕ_0 =90° (dotted line)

Although we expected the exponential behaviour (6.10), the actual error is such that the higher is the spatial bandwidth of the incident field in the points of the profile with a minimum curvature radius, say $\widetilde{w}$, the higher is the error. In fact, the worst case is when the propagation is parallel to the minor axis, since $\widetilde{w} \approx \beta$. On the contrary, the error decays more rapidly when ϕ_0=0, being almost zero the spatial bandwidth at the ellipse ends.

6.3. Application of the optimal positioning criterion

As seen in the Section 5.3, when the collocation points are uniformly distributed assuming a value of χ slightly larger than unity, the scattered field is sampled at a frequency slightly larger than β, thus resulting in possible undersampling on some portions of the observation curve. In order to clarify this point, the qualitative behaviour of the spatial bandwidth of the field due to an incident plane wave on the elliptic profile is shown in the figure 6.15. The envelope of the spectral content of all the possible incident fields gives the curve of the function w_ε, representing the local sampling bandwidth that provides a constant error, that varies along the curvilinear abscissa s according with the profile's curvature.

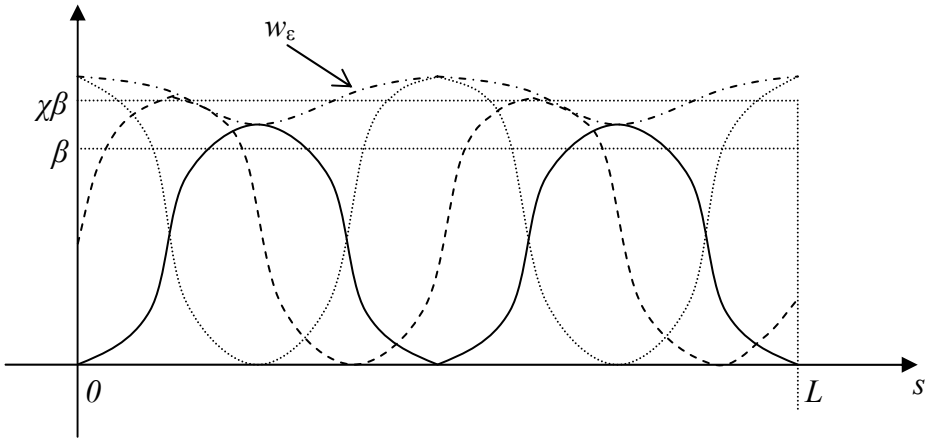


Figure 6.15 – Qualitative behaviour of the actual bandwidth of the incident field on the elliptical boundary, for the following conditions of incident plane wave: $\phi_0 = 0$ (solid line), $\phi_0 = 45°$ (dashed line), $\phi_0 = 90°$ (dotted line). The dash-dotted line indicates the function w_ε

In the previous chapter it has been shown that, taking into account the profile curvature, an "optimal" curve parameterization is found, that ensures an uniform band-limitation error on the observation domain. The optimal parameterization thickens the testing point in the region with the higher values of w_ε. In the case of an elliptic profile the analytic expression for the curve's parameterization is given by:

$$\frac{n}{N} = \frac{s}{\chi L} + \left(1 - \frac{1}{\chi}\right)\frac{F(\theta, e)}{4K(e)} \qquad (6.16)$$

where s is the curvilinear abscissa on the profile, $K(e)$ is the complete elliptic integral of the first kind, $F(\theta,e)$ is the incomplete elliptic integral of the first kind, and $e^2 = 1 - (a^2/b^2)$.

The expression (6.16) gives the optimal positions of the N collocation points for n assuming values in [1,N]. Since the elliptic integral cannot be inverted analytically, the right end of the (6.16) is evaluated in a large set of positions on the boundary, thus building a look-up table, then the solution of the equation for each value of n is obtained by conveniently interpolating the table data.

This positioning criterion has been applied to the above described MAS implementation in the case of the elliptic scattering cylinder. The first case involves an ellipse of semi-axes (10λ, 6λ) under the incidence of two plane waves propagating along the axis directions. The results in figure 6.16 show that the field evaluation error, obtained by the introduction of the optimized positioning criterion in the MAS implementation, fits the exponential behaviour of the band-limitation error (6.10), thus increasing the accuracy with the respect to the standard uniform test point positioning case.

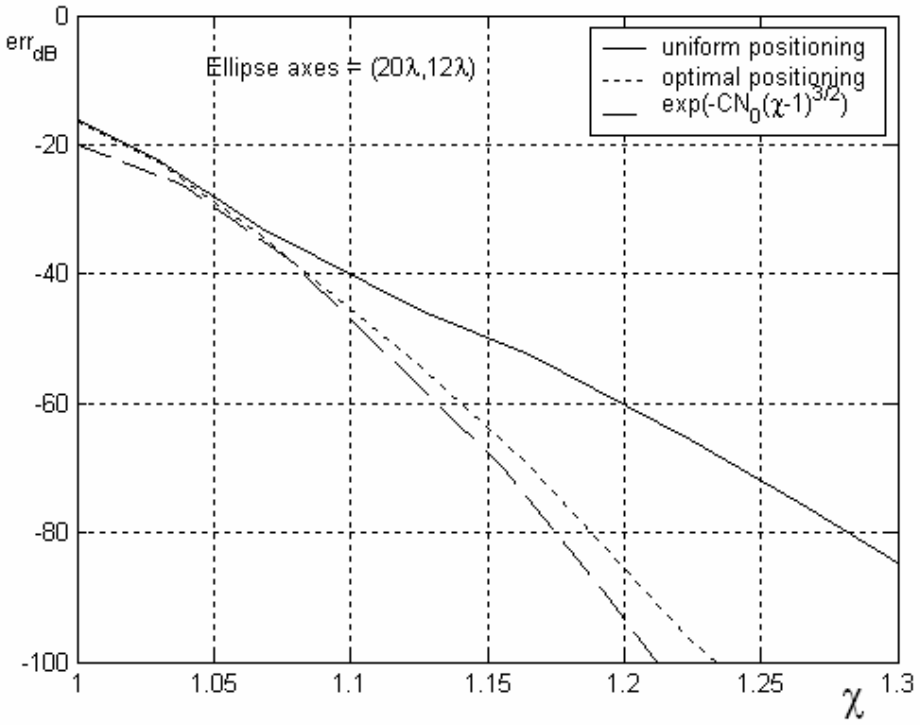


Figure 6.16 – Scattered field evaluation error for an ellipse of semi-axes (10λ, 6λ). The solid line indicates the error of the MAS with uniform test point positioning, the dotted line indicates the MAS error with the proposed optimized positioning, and the dashed line shows the theoretic band-limitation error behaviour.

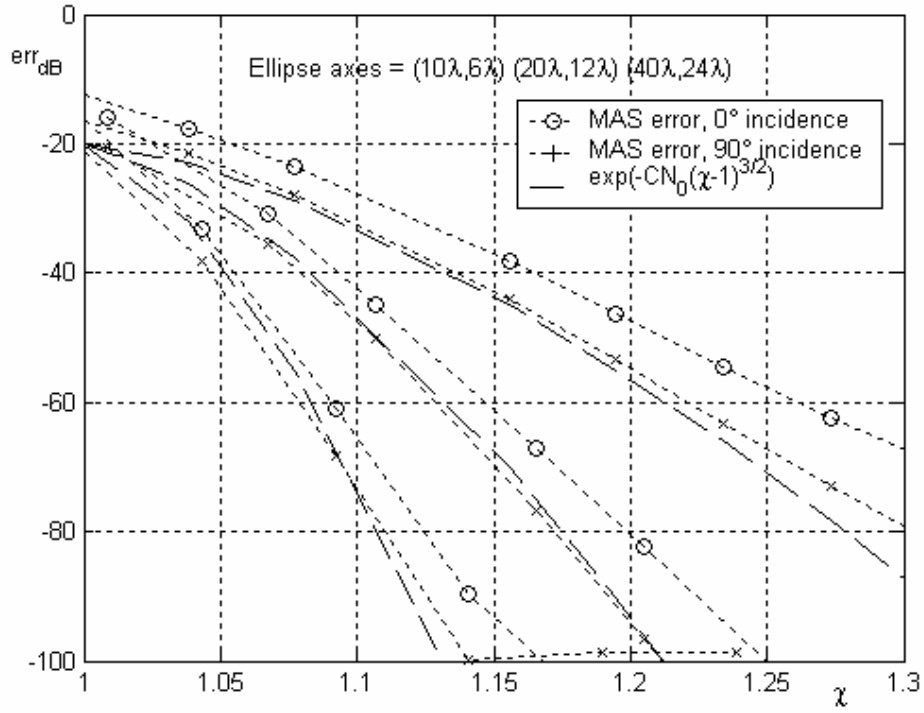


Figure 6.17 - Comparison between the scattered field evaluated by the MAS with the optimized positioning criterion (dotted lines) and the theoretich band-limitation error (dashed lines), for various sizes of the scattering ellipses and incident plane waves.

Moreover, three scattering ellipses of semi-axes (5λ, 3λ), (10λ, 6λ) and (20λ, 12λ) illuminated by two plane waves propagating along the major and minor axis, respectively, have been considered. The figure 6.17 shows the behaviour of the error in the evaluation of the scattered field in the six considered cases. A comparison with the asymptotical trend of the band-limitation error is performed.

The comparison confirms that the use of a proper curve parameterization allows to obtain an error behaviour that complies

with the exponential decay of the band-limitation error, that can be estimated in advance by means of the only knowledge of the scattering geometry.

6.4. The FMM-MAS

To improve the computational performances, an FMM approach has been applied to the impedance matrix inversion of the proposed MAS-based technique, as formulated in the chapter IV.
It must be noted that, in the FMM based technique the whole impedance matrix is never explicitly computed, because its role is played by the M-by-M matrices $T_{\hat{k}_i}$ (representing the aggregate interactions between clusters) together with the excitation vectors $V_{\hat{k}_i}$.
As previously discussed, the memory needed to store these data is proportional to $N^{3/2}$, while a dense N-by-N matrix would requires the storage of N^2 values. On the other hand, the matrix inversion is performed with a Generalized Minimal Residual (GMRES) based iterative process which essentially requires an (impedance) matrix-by-(current) vector product at each step. As seen in chapter VI, the FMM technique allows to evaluate this product with a complexity $O(N^{3/2})$, instead of the classical $O(N^2)$.
The numerical analysis has been performed by considering the scattering of a plane wave from several ellipses of increasing size. The couple made by the major and minor axes varies from a minimum corresponding to (150λ, 90λ) to a maximum corresponding to (400λ, 240λ). A MAS technique capable of evaluating the scattering by these geometries with a minimal number of unknowns has been implemented..
In the GMRES-based iterative method the FMM parameters have been set to $L=1.1kd_{max}$ and $N_k=2L$, according to the usual prescriptions [62].
Figure 6.18 shows a comparison between the memory required to store the full impedance matrix (needed in classical MoM and MAS)

and the memory requirement in the FMM-MAS implementation. As previous discussed, the use of a low-redundancy representation for the impedance matrix results in a significant memory saving. Figure 6.19 reports the computational time required to compute the matrix-by-vector product, in both the standard and FMM cases, as a function of the size of the scatterer. Even if for small problem sizes the complications introduced with the FMM is not compensated by the FMM speed up, we note that the curves relative to fast case grow with a smaller order with the number of unknowns, this assuring an advantage for large problems [71].

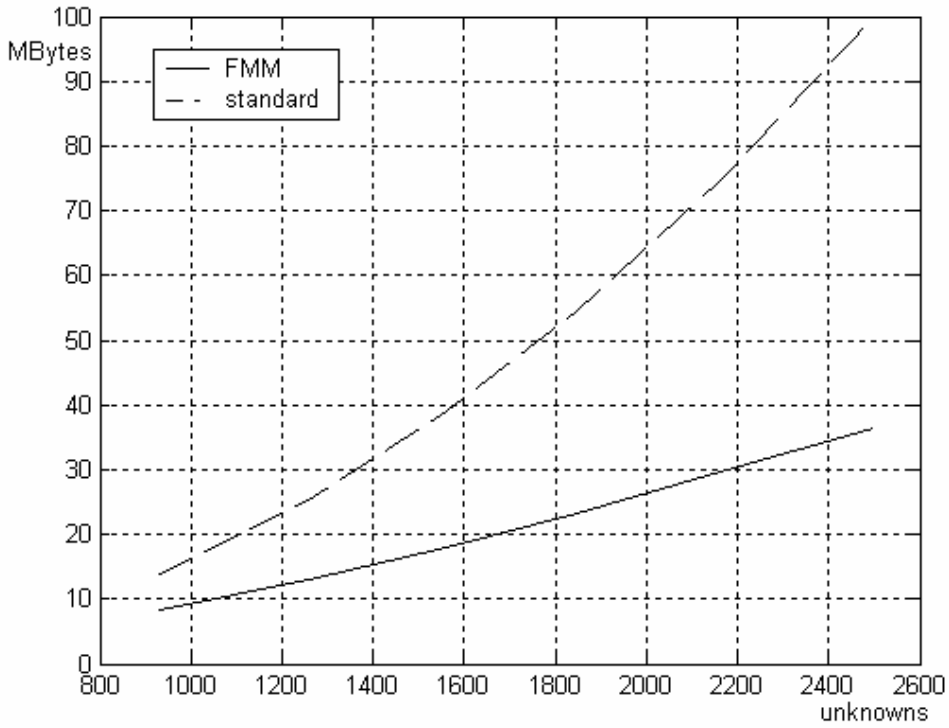


Figure 6.18 – Memory needed to store the classical impedance matrix (dashed line) compared to the storage of the low-redundancy representation exploiting the FMM (solid line)

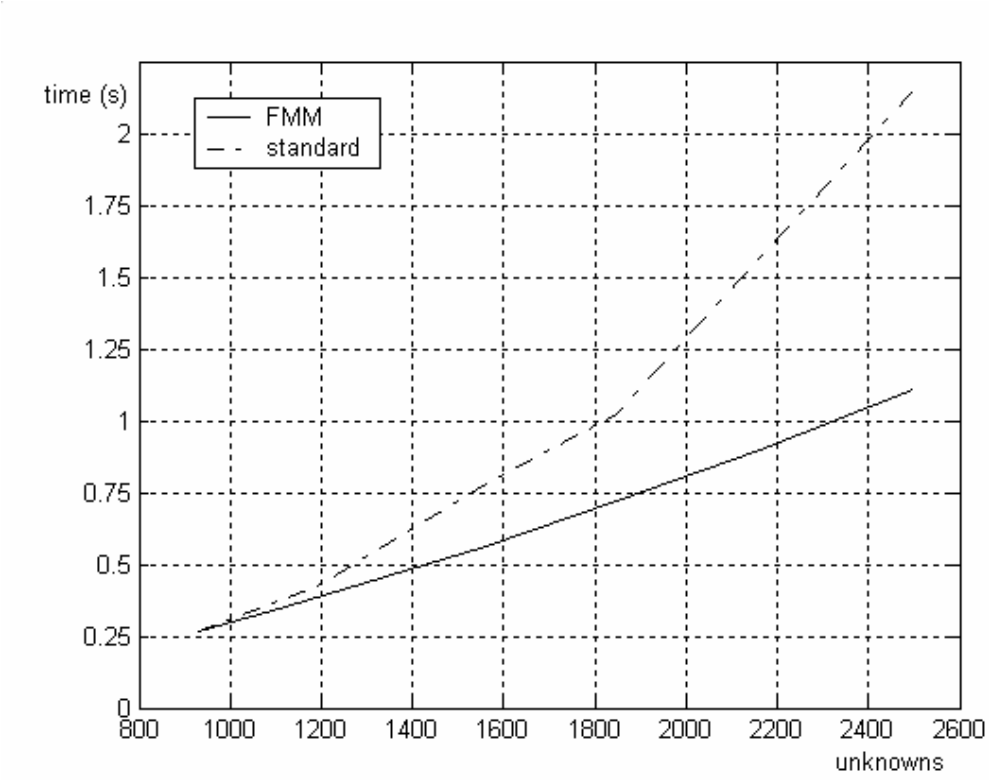


Figure 6.19 – Comparison between the computational time needed to compute the classical product between impedance matrix and current vector in the standard (dashed line) and in the FMM schemes (solid line) by a PC equipped with a Pentium 4 CPU at 3.20GHz and 2 GBytes of RAM.

The overall error, ε, of the scattered field evaluated by means of the described technique is separable into three components: the error ε_{mas} due to the MAS implementation of the problem; the error ε_{iter}, due to the iterative solution process, with respect to the solution by direct

inversion; the error, ε_{fmm}, introduced by the usage of the FMM scheme, which approximates the matrix-by-vector products needed in the iterative solution.
These last error terms are numerically computed for the considered geometries. In particular, it is found ε_{iter}<-75dB and ε_{fmm}<-66dB, while the overall error reaches the maximum value of -65dB. The consideration made in the previous sections ensures that for the chosen oversampling rate (χ=1.2) and scatterer size the error ε_{mas} is negligible with respect to the total error. As a check, a reference solution has been evaluated by using a significantly larger oversampling ratio (χ=1.6), obtaining ε_{mas}<-100dB.
It is important to note that the value of ε_{fmm} can be reduced, if needed, by using a larger value for L and/or a higher number N_k of directions when evaluating the factorized expression of the Green's function [71].

6.4.1 Preconditioning

As pointed out in the chapter 4, another relevant point related to the computational complexity of the problem involves the choice of an appropriate preconditioning technique that allows to reduce the number of iterations required to reach the GMRES convergence. The first choice was to adopt a block Jacobi preconditioning, but the results were pour, as shown in figure 6.20. Then, we applied the preconditioning technique described in the chapter IV and based on the matrix of the MAS "near interaction", which is explicitly evaluated and available: this approach allowed to strongly reduce the number of interactions required to reach the convergence, by providing an exponential-like decay for the residual sequence (see figure 6.20). The entity of the iteration-saving factor depends on the precision required (i.e. the value of the residual chosen to stop the process).

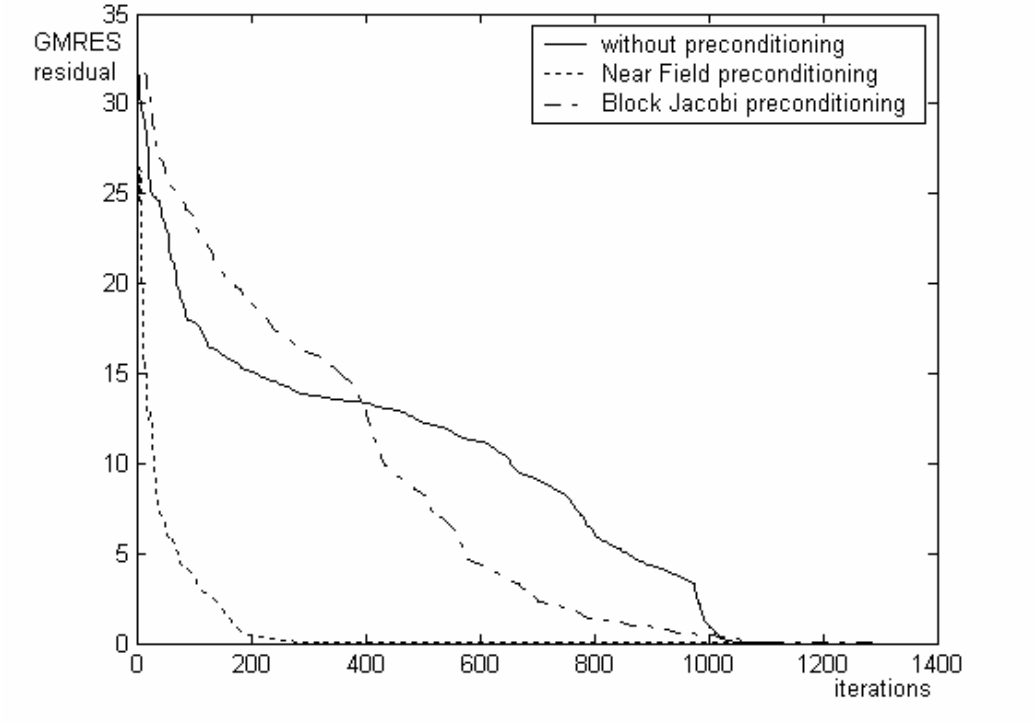


Figure 6.20 – Sequences of GMRES residuals. Solid line: no preconditioner used. Dotted line: preconditioning based on the MAS near interaction data. Dashed line: Block Jacobi preconditioning.

6.5. The MAS for 3D scatterers

The result of the application of the MAS algorithm to the analysis of three-dimensional PEC scatterers are presented in this Section. The numerical efficiency is of crucial interest in the evaluation of the RCS of objects of relevant size. The analysis of scatterers of relevant size (for example, the evaluation of the RCS) is usually approached by exploiting asymptotical methods, even if they provide a reduced precision with respect to the integral methods. In fact, the use of full-wave technique like the standard MoM, requires a prohibitive computational effort that becomes soon unsustainable as the size of the problem exceeds a few wavelengths.
In order to show the potential of the MAS in reducing the size of the problem, even for 3D problems, the RCS of some 3D geometries have been evaluated by means of our MAS implementation and, for the sake of comparison, with a standard MoM algorithm featuring RWG basis functions.
First, we considered a simple, canonical geometry: the sphere. This choice will allow us to evaluate the precision of both the MoM and the MAS results, since, in the spherical case, the value of the scattered field is analytically known. With reference to a sphere with the radius equal to 3λ, the mesh of the MoM algorithm, featuring 5154 RWG basis function, and the displacement of the MAS sources and testing points, involving 1136 impulsive functions are shown under Figures 6.21 and 6.22, respectively. Despite the lower number and the simpler kind of basis function used, the RCS evaluated with the MAS algorithm is more precise than the one provided by the MoM, the MAS featuring a mean square error smaller than -21dB, the MoM reaching an error -16dB (see figure 6.23).

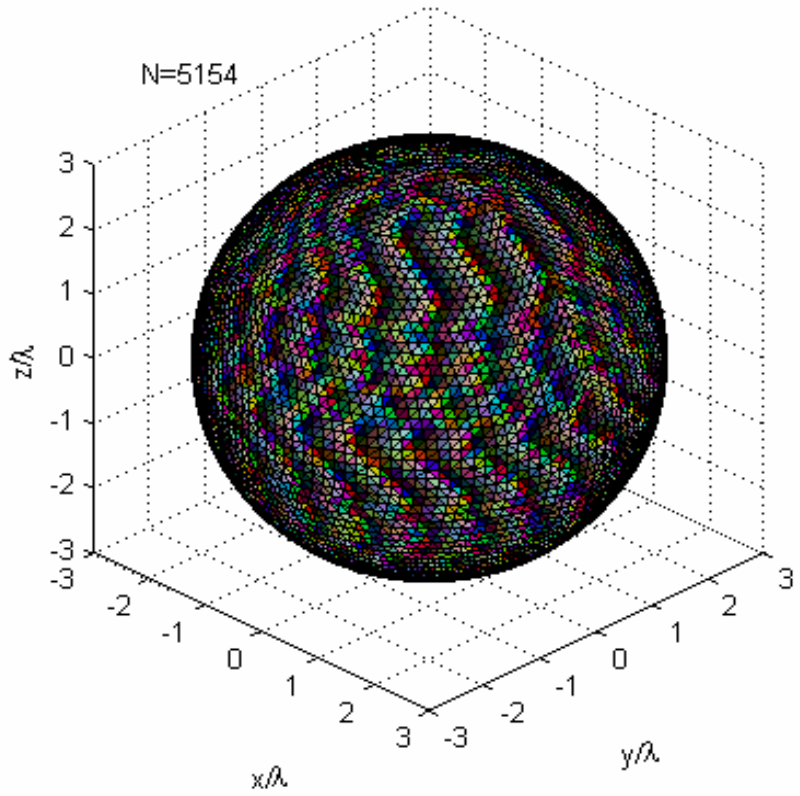


Figure 6.21 – MoM mesh featuring RWG basis function, spherical case

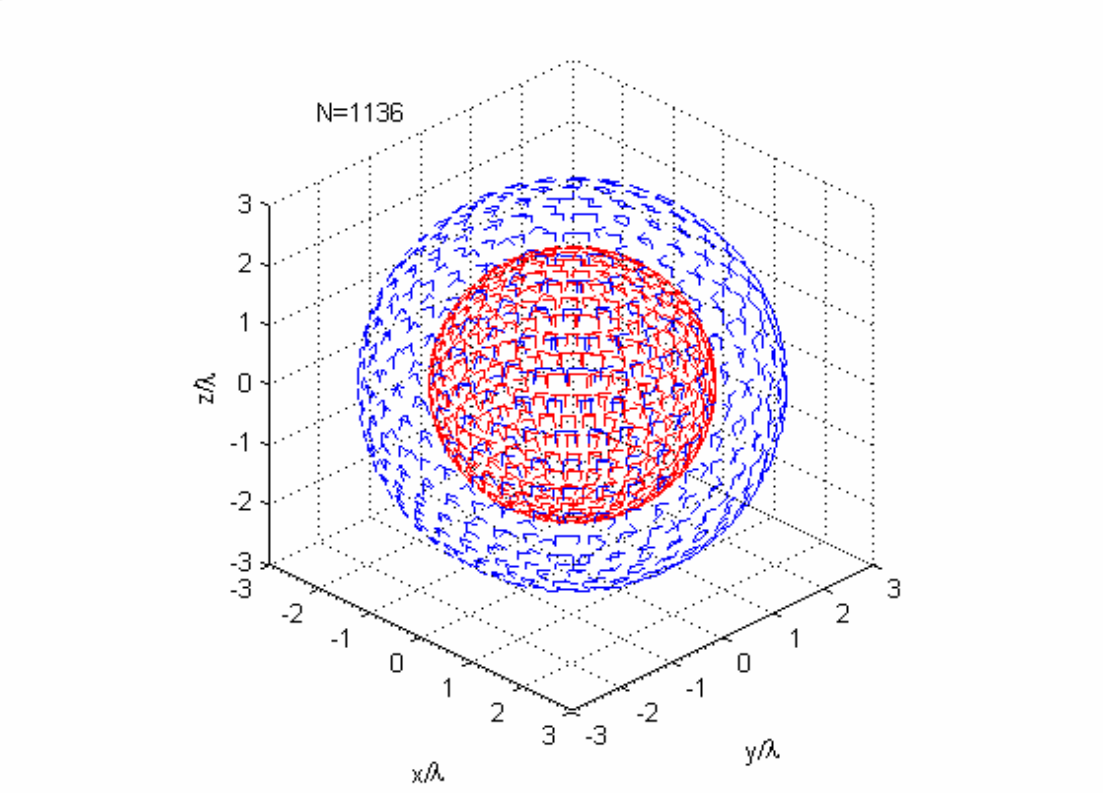


Figure 6.22 – The spherical case studied by means of MAS sources and impulsive testing functions

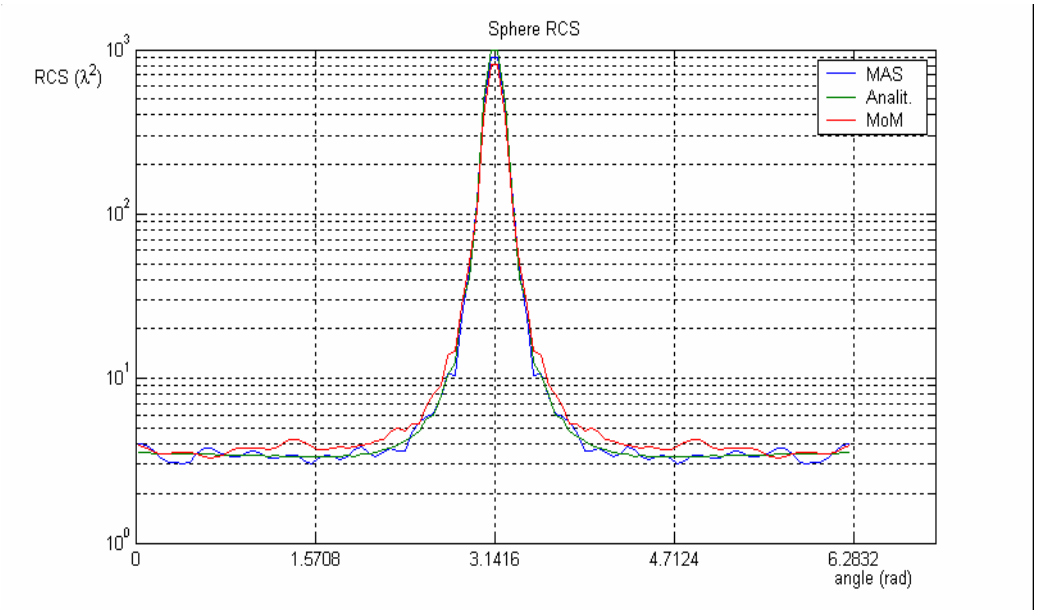


Figure 6.23 – Comparison between the RCS compared by the MAS and the MoM with the analytical solution, spherical case

A more realistic scattering geometry we considered is made by the superimposition of an upper hemisphere, a circular cylinder and a lower hemisphere. In the following, we will refer to this geometry as the "pill".
The mesh of the MoM algorithm, featuring 4528 RWG basis function, and the displacement of the MAS sources and testing points, involving 1252 impulsive functions are shown in the figures 6.24 and 6.25, respectively.
The results of the RCS evaluations obtained by means of the two methods (MAS and MoM) are shown in figure 6.26. The two results differ for -28dB in mean square value. This difference is comparable with the precision found for the RCS evaluation found in the spherical case, where the known analytic result has been used as reference.
This means that the two evaluations substantially have the same precision, even if the MAS requires a reduced computational effort.

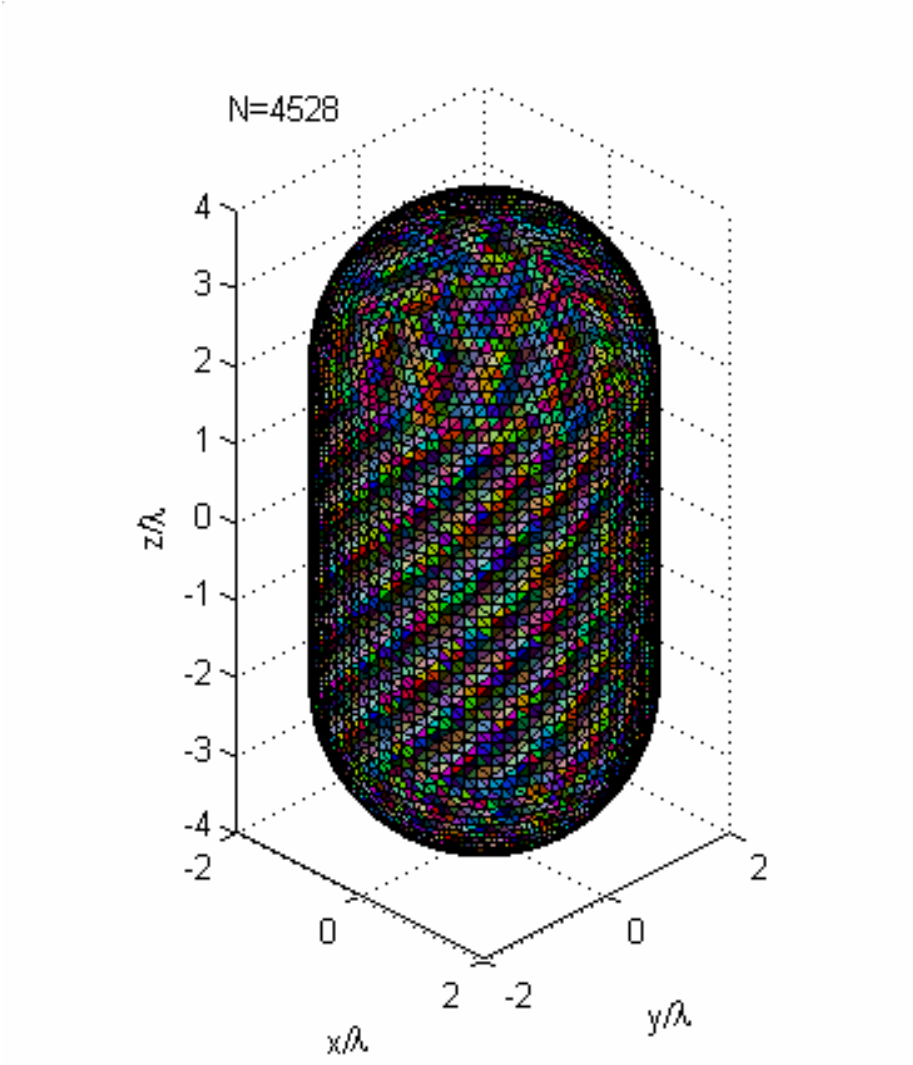


Figure 6.24 – MoM mesh featuring RWG basis function, "pill" case

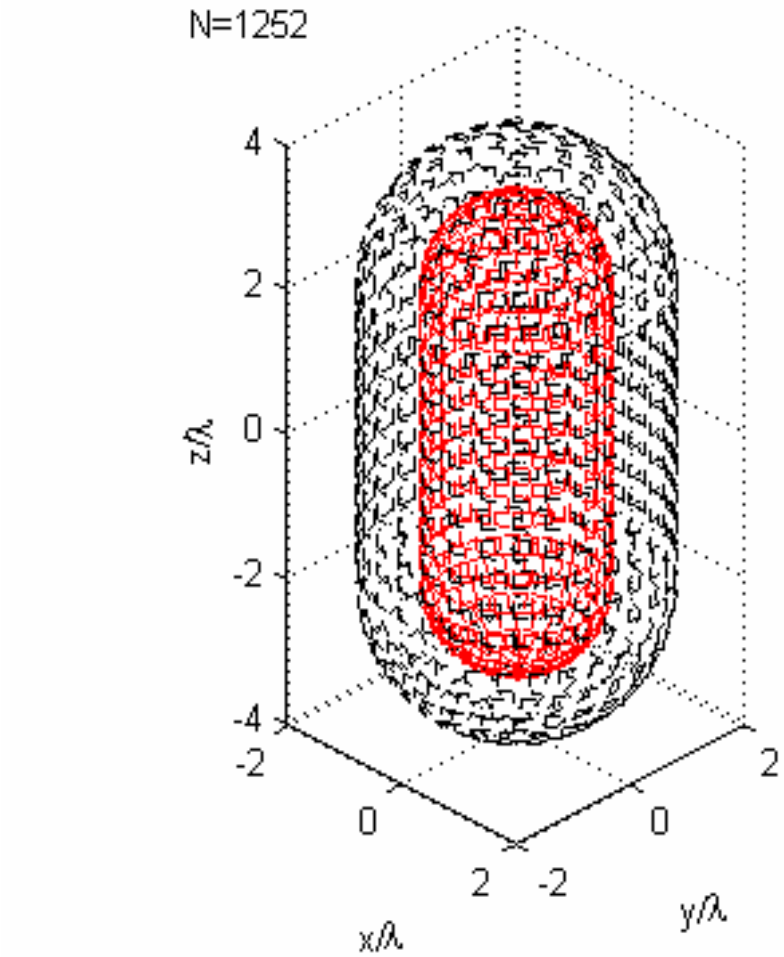


Figure 6.25 – The "pill" case studied by means of MAS sources and impulsive testing functions

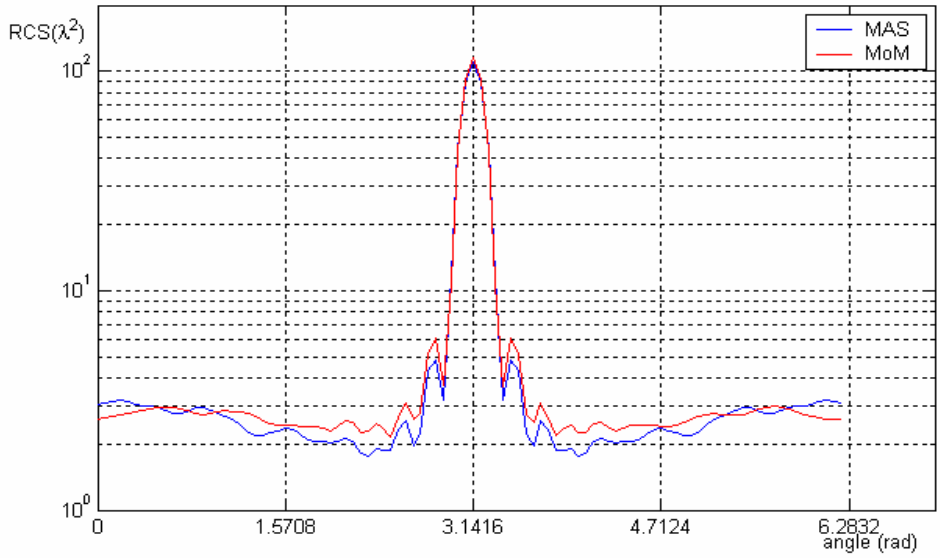


Figure 6.26 – Comparison between the RCS compared by the MAS and the MoM with the analytical solution, "pill" case

To introduce a further realistic detail, we considered a 3D geometry with an edge. The scatterer, that we will refer to as the "bullet", is build by laying an hemisphere upon a circular cylinder. Since the scattering surface locally presents a non-finite curvature, the local band-limitation hypothesis cannot apply.
To manage this case we must take into account that the local spatial bandwidth is expected to rise, so that we have to:

- thicken the testing point in the region nearby the edge, in order to better enforce the boundary condition
- place the auxiliary sources closer to the scattering edge, so that they can produce a field not band-limited to β anymore

In this way, the auxiliary sources distribution is no more conformal to the scattering geometry. The mesh of the MoM algorithm, featuring 3987 RWG basis function, and the displacement of the MAS sources

and testing points, involving 1238 impulsive functions are shown in the figures 6.27 and 6.28.
As seen in figure 6.29, the RCS evaluation made by the MAS and the MoM algorithms, differing for -24dB, can be considered to have the same precision, for the same reasoning followed in the “pill” case.

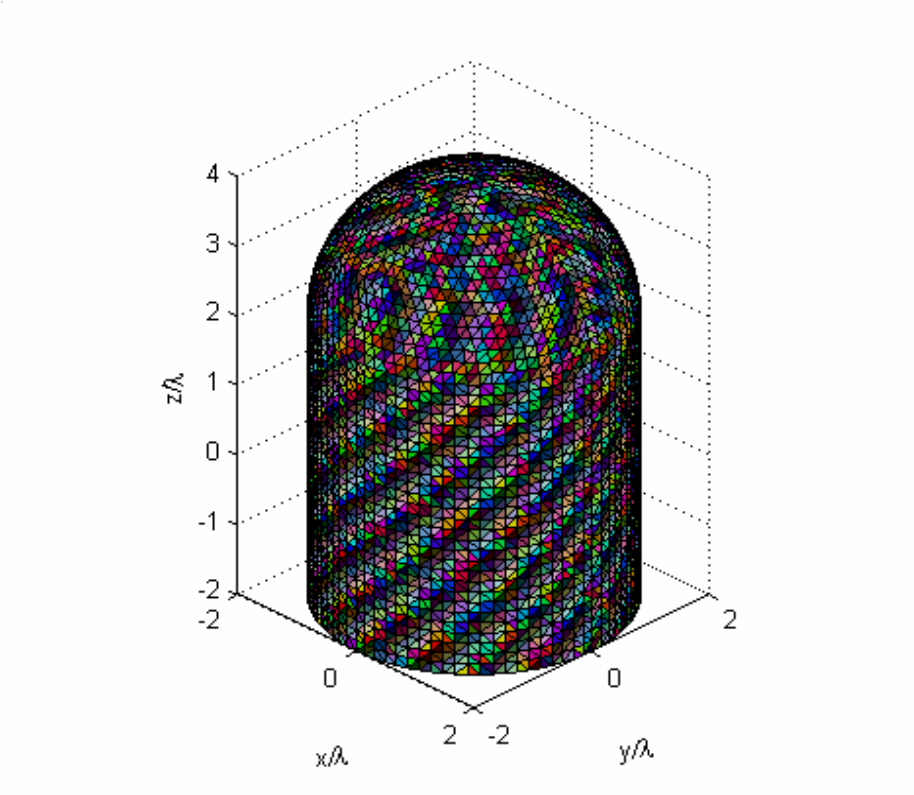
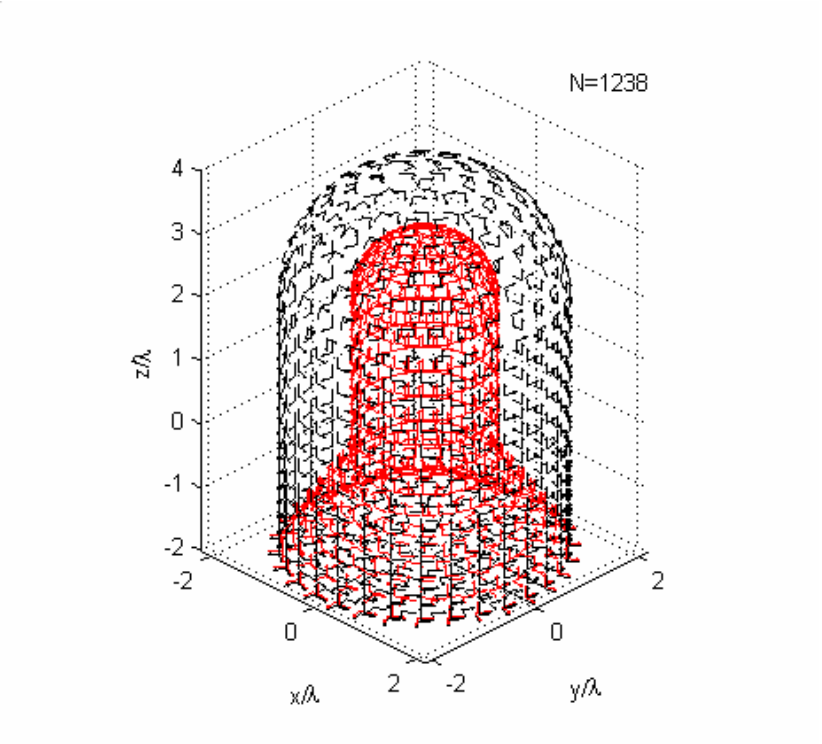


Figure 6.27 – MoM mesh featuring RWG basis function, “bullet” case


N=1238
z/λ
x/λ
y/λ

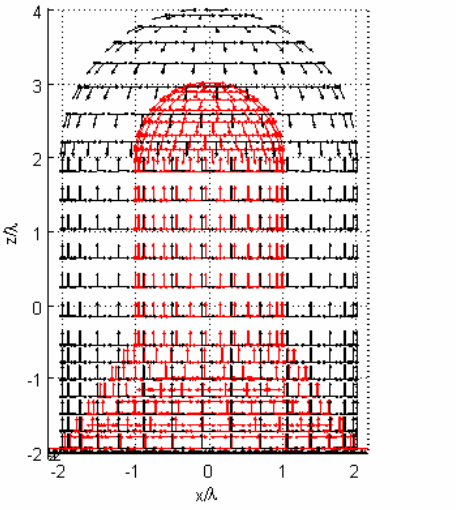

z/λ
x/λ

Figure 6.28 – The "bullet" case studied by means of MAS sources and impulsive testing functions. The auxiliary surface is not exactly conformal to the scattering profile

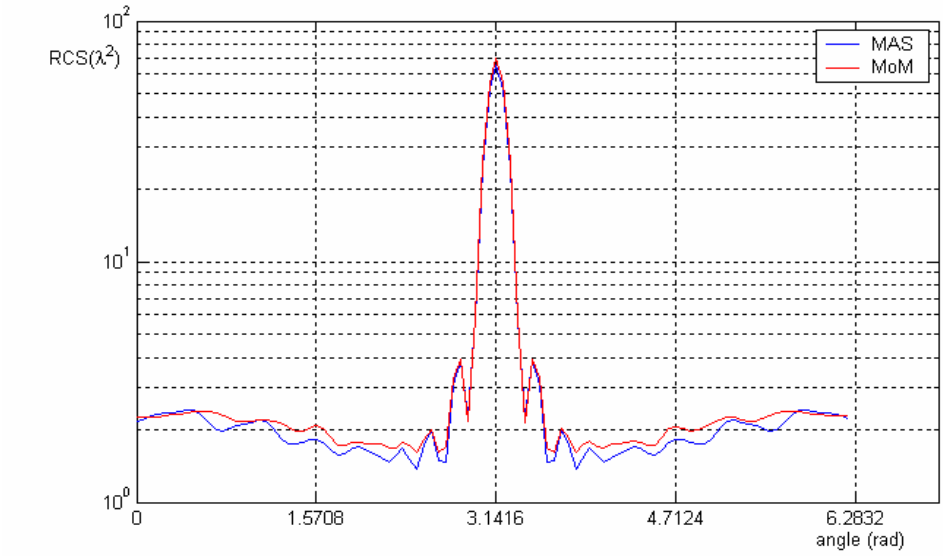


Figure 6.29 – Comparison between the RCS compared by the MAS and the MoM with the analytical solution, "bullet" case

6.6. Conclusions

In conclusion, we introduced a criterion for the positioning of the auxiliary sources and the testing points in the MAS approach thus allowing the evaluation of the field scattered by a PEC scatterer by means of a practically minimal number of unknowns (slightly larger than 2 per wavelength.

Moreover, we have verified that the error related to the evaluation of the scattered field agrees with the band-limitation error, that is theoretically known and depends on the oversampling ratio (i.e. the ratio between the number of unknown used and the number of the

degrees of freedom of the field). Accordingly, this ratio can be chosen in advance according to the desired accuracy.
Moreover, it has been shown that the Fast Multipole paradigm can be applied to the MAS in order to obtain a relevant reduction of the computational costs related to the evaluation, storage and inversion of the impedance matrix.
Finally, it has been shown that a MAS algorithm can be used to evaluate the field scattered by a three-dimensional PEC object, even in presence of edges, providing a precision comparable to a classical MoM implementation, with a strongly reduced number of unknowns.

References

[1] O.M. Bucci, "Computational Complexity in the solution of large antenna and scattering problems", *Radio Science*, Vol.40, 2005

[2] O.M.Bucci, G. Franceschetti, "On the degrees of freedom of scattered fields", IEEE trans. Antennas Propagat., vol. 37, pp. 918-926, July 1989.

[3] O. M. Bucci , C. Gennarelli, C. Savarese, "Representation of Electromagnetic Fields over Arbitrary Surfaces by a Finite and Nonredundant Number of Samples", *IEEE Trans. Antennas. Propagat., Vol 46, March 1998*

[4] O.M. Bucci, G. Di Massa, "Use of characteristic modes in multiple-scattering problems, *J. Phys. D: Appl. Phys.*, vol. 28, No. 11, 1995.

[5] Yousef Saad. *Iterative Methods for Sparse Linear Systems,* Pws Publishing Co., 1996

[6] R. Barrett and M. Berry and T. F. Chan and J. Demmel and J. Donato and J. Dongarra and V. Eijkhout and R. Pozo and C. Romine and H. Van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM 1994

[7] W.C. Chew, J.M. Jin, C.C. Lu, E. Michielssen, J. M. Song, "Fast Solution Methods in Electromagnetics", *Trans. Antennas Propag*., Vol. 45, No. 3, March 1997.

[8] A. Boag, R. Mittra, “Complex Multipole Beam Approach to Electromagnetic Scattering Problems”, *IEEE Trans. Ant. Prop.*, vol. 42, March 1994.

[9] V.V.S. Prakash, R.Mittra, “Characteristic Basis Function Method: A new technique for fast solution of integral equations”, Microw. and Opt. Tech. Letters, Jan. 2003.

[10] L. Matekovits, G. Vecchi, P. Pirinoli, G. Dassano, M. Orefice, ”Reduced-complexity MoM simulation of printed structures”, *Digest of 2000 AP-S Int’l Symp.*, pp. 352 - 355, July 2000, Salt Lake City, Utah, USA.

[11] L. Matekovits, G. Vecchi, G. Dassano, M. Orefice, ”Synthetic function analysis of large printed structures: the solution space sampling approach”, *Dig. of 2001 IEEE Antennas and Propagation Society Int’l Symp.*, pp. 568 - 571, Jul. 2001, Boston, Massachusetts.

[12] L. Matekovits, G. Vecchi, P. Pirinoli, G. Dassano, F. Vipiana, M. Orefice “Discretization-diversity synthetic-function technique for the simulation of large printed antennas”, *XIV RiNEM*, Ancona, Italy, 2002

[13] L. Matekovits, G. Vecchi, V.A. Laza, “Degrees of Freedom And Synthetic Functions In The Analysis of Large Antennas”, *Digest of the 2004 URSI Int. Symp. On Electromagnetic Theory*, Pisa, Italy, May 2004.

[14] V.D. Kupradze, M.A. Aleksidze, “The method of functional equations for approximate solution of some boundary problems”, J. of Appl. Math. And Math. Physics, 4, 1964.

[15] I.N. Vekua, *The new methods for solving elliptic equations*, Jonh Wiley, New Jork 1967

[16] M.A. Aleksidze, *Solution of boundary problems by expansion into a nonhortogonal series*, Nauka, Moskow, 1978.

[17] R. Popovidi-Zaridze, D. Karkashadze, G. Ahvlediani, and J. Khatiashvili,"Investigation of possibilities of the method of auxiliary sources in solution of two-dimensional electrodynamics problems," *Radiotech. Electron.*, vol. 22, no. 2, 1978.

[18] D. Karkashadze and R. Zaridze, "The Method of Auxiliary Sources in Applied Electrodynamics", *Latsis Symposium on Computational Electromagnetics,* Zurich, 19-21 September, 1995, Invited Lectures. pp. 163-180.

[19] F. G. Bogdanov, D. D. Karkashadze, and R. S. Zaridze, "The method of auxiliary sources in electromagnetic scattering problems," *Generalized Multipole Techniques for Electromagnetic and Light Scattering*, 1999

[20] O.M. Bucci, G. D'Elia, M. Santojanni, "Non Redundant Implementation of the Method of Auxiliary Sources for 2-D Smooth Geometries", IEE Electronics Letters, Vol. 41 Issue 22, October 2005

[21] F.X. Canning, "The impedance matrix localization (IML) method for moment-method calculations", Antennas and Prop. Mag, Volume 32, Issue 5, Oct. 1990 Page(s):18 – 30

[22] F.X. Canning, "Interaction Matrix Localization (IML) permits Solution of Larger Scattering Problems", *IEEE Trans. on Magnetics*, vol.27, no.5, Sept. 1991.

[23] F.X. Canning, "Improved Impedance Localization Method", *IEEE Trans. on Ant. and Prop*., vol. 41, no. 5, May 1993

[24] B. Z. Steinberg and Y. Leviatan, «On the use of wavelet expansions in the method of moment», IEEE Trans. A. P., Vol. 41, No. 5, Mai 1993

[25] P. Pirinoli, G. Vecchi and L. Matekovits, «Multiresolution analysis of printed antennas and circuits: a dual-isoscalar approach», IEEE Trans. A. P., Vol. 49, No. 6, June 2001

[26] R. Loison, R. Gillard, J. Citerne, "A Multiresolution MoM Analysis of Multiport Structures Using Matched Terminations", *IEEE Trans. on Microw. Theory and Tech.*, vol.49,no.1, January 2001

[27] R.L. Wagner, W.C. Chew, "A Study of Wavelets for the Solution of Electromagnetic Integral Equations", *IEEE Trans. on Antennas and Prop*., vol 43, no.8, August 1995.

[28] Z. Xiang, Y. Lu, "An Effective Matrix Transform Approach for Efficient Solutions of Electromagnetic Integral Eqautions", *IEEE Trans. on Antennas and Prop*., vol 45, no.8, August 1997

[29] W.L. Golik, "Wavelet Packets for Fast Solution of Electromagnetic Integral Equations", *IEEE Trans. on Antennas and Prop*., vol 46, no. 5, May 1998.

[30] J. H. Lin and W. C. Chew, "BiCG-FFT T-Matrix Method for Solving for the Scattering Solution from Inhomogeneous Bodies", *IEEE Trans. on Microw. Theory and Tech.*, Vol. 44, no. I , July 1996

[31] Y. Zhuang, Ke-Li Wu, C. Wu, J. Litva, "A Combined Full-Wave CG-FFT Method for Rigorous Analysis of Large Microstrip Antenna Arrays", *IEEE Trans. Ant. Prop.*, vol44, no.1, Jan. 1996.

[32] W.B. Lu, T.J. Cue, X.X. Yin, Z. G. Quian, W. Hong, "Fast Algorithms for Large-Scale Periodic Strucures Using Subentire

Domain Basis Functions", *IEEE Trans. Ant. Prop.*, vol.53, no.3, March 2005.

[33] N. Engheta, W. D. Murphy, V. Rokhlin, M. S. Vassiliou, "The Fast Multipole Method for Electromagnetic Scattering Problems", *IEEE Trans. Antennas. Propagat.*, June 1992.

[34] Rao, S.M., D.R. Wilton, A.W. Glisson, "Electromagnetic scattering from surfaces of arbitrary shape",*IEEE Trans. Ant. and Prop*., vol. 30, pp. 409–418, May 1982.

[35] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 11, pp. 674-693, July 1989.

[36] L. Greengard, V. Rokhlin, "A Fast Algorithm for Particle Simulations", *J. Comp. Physics*, 1987

[37] V. Rokhlin, "Rapid Solution of Integral Equations of Classical Potential Theory", *J. Comp. Physics* 60, 187-207, (1985)

[38] V. Rokhlin, "Rapid Solution of Integral Equations of Scattering in Two Dimensions", *J. Comp. Physics* 86, 414-439, (1990)

[39] C. C. Lu and W. C. Chew, "A multilevel algorithm for solving boundary-value scattering," *Microwave Opt. Tech. Lett*., vol. 7, no. 10, pp. 466–470, July 1994.

[40] J. M. Song and W. C. Chew, "Multilevel fast-multipole algorithm for solving combined field integral equations of electromagnetic scattering," *Microwave Opt. Tech. Lett*., vol. 10, no. 1, pp. 14–19, Sept. 1995.

[41] Y. Leviatan, A. Boag, and A. Boag, "Generalized formulations for electromagnetic scattering from perfectly conducting and homogeneous material bodies—Theory and numerical solution,"

IEEE Trans. Antennas Propagat., vol. 36, no. 12, pp. 1722–1734, Dec. 1988.

[42] Y. Leviatan, "Analytic continuation consideration when using generalized formulation for scattering problem", *IEEE Trans. Antennas Propagat.*, vol. 38, issue 8, Aug. 1990.

[43] R. F. Harrington, *Field Computation by Moment Methods*. New York: Macmillan, 1968.

[44] T.K.Sarkar, A.R.Djordjevic, B.K.Kolundzija, 'Method of Moments Applied to Antennas', chapter in *The Handbook of Antennas in Wireless Communications*, Ed. Lal Godara, CRC Press, Boca Raton, Florida, 2001.

[45] S. Eisler and Y. Leviatan, "Analysis of electromagnetic scattering from metallic and penetrable cylinders with edges using a multifilament current model," *Proc. Inst. Elect. Eng.* , pt. H, vol. 136, pp. 431–438, Dec. 1989.

[46] Anastassiu, H.T.; Kaklamani, D.I.; Economou, D.P.; Breinbjerg, O.; "Electromagnetic scattering analysis of coated conductors with edges using the method of auxiliary sources (MAS) in conjunction with the standard impedance boundary condition (SIBC)", *IEEE Trans. Antennas. Propagat.,* Vol. 50, Issue 1, Jan. 2002

[47] Hristos T. Anastassiu, Dimistrios G. Lymperopoulos, Dimitra I.Kaklamani, "Accuracy Analysis and Optimization of the Method of Auxiliary Sources (MAS) for Scattering by a Circular Cylinder", *IEEE Trans. Antennas Propagat.*, vol 52, No. 6, June 2004

[48] Dimitra I.Kaklamani, Hristos T. Anastassiu, "Aspects of the Method of Auxilairy Sources (MAS) in Computational Electromagnetics", *IEEE Antenna's and Propagation Magazine,* Vol. 44, No. 3, June 2002

[49] Georgios K. Advikos, Hristos T. Anastassiu, "Computational Cost Estimations and comparison for Three Methods of Applied Electromagnetics (MoM, MAS, MMAS)", *IEEE Antennas Propagat. Magazine*., vol. 47, no.1, 2005

[50] O.M.Bucci, G. Franceschetti, "On the spatial bandwidth of scattered fields", IEEE trans. Antennas Propagat., vol. 35, pp. 1445-1455, Dec. 1987.

[51] O.M.Bucci, G. D'Elia, "Advanced Sampling Techniques in Electromagnetics", *Review of Radio Science 1993-1995*, Oxford University Press, London, 177-204, 1996.

[52] J.J Dongarra, F. Sullivan, "The top 10 algorithms", *Comput. Sci. Eng*., 2:22-23, 2000

[53] Y.Saad and M.H.Schultz, "GMRES: A Generalized Minimal Residual algorithm for solving nonsymmetric linear systems", *SIAM Sci. Stat. Comput.*, 7 (1986), pp. 856-869

[54] I. Ipsen, C. Meyer, "The idea behind Krylov methods", *Amer. Math. Monthly,* vol 105, n.10, 1998

[55] M. Hestenes, E. Stiefel, "Methods of conjugate gradients for solving linear systems" *J. Res. Nat. Bur. Stand*. 49 (1952)

[56] R. Fletcher, *Conjugate gradient methods for indefinite systems, in Numerical Analysis Dundee*, G. Watson ed., Berlin, New York 1976, Springer Verlag.

[57] O. Widlund, "A Lanczos method for a class of non symmetric systems of linear equations", *SIAM J. Numer. Anal*., 15 (1978)

[58] W. E. Arnoldi, "The principles of minimized iteration in the solution of the matrix eigenvalue problem", *Quart. Appl. Math*., 9 (1951)

[59] J. Song, C. C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propagat.*, vol. 45, no. 10, pp. 1488-1493, Oct 1997.

[60] K. Sertel, J. L. Volakis, "Incomplete LU preconditioner for FMM implementation", *Microw. Opt. Tech. Lett*., vol. 26, n.4, Aug. 2000

[61] M. Abramowitz, I. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1965

[62] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: a pedestrian prescription,"IEEE *Ant. Propag. Mag.*,vol. 35, no. 3, pp. 7–12, June 1993.

[63] P. Morse, H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill, 1953

[64] O.M.Bucci, C.Gennarelli, G.Riccio, C.Savarese, "Fast and accurate far field evaluation from a non redundant, finite number of plane-polar measurements", *1994 IEEE Antennas and Propagation Int. Symposium Digest*, June 1994, Seatte, USA

[65] O.M.Bucci, C.Gennarelli, G.Riccio, C.Savarese, "Precise interpolation of electromagnetic fields over a cylinder from a non redundant, finite number of samples", *Proceedings of the Progress in Electromagnetics Research Symposium*, July 1994, Nordwijk, NL

[66] O.M.Bucci, C.Gennarelli, C.Savarese, "Non-redundant representation of electromagnetic fields", *Proceedings of the journees Int. de Nice sur les Antennes,* November 1994, Nice, France

[67] O. M. Bucci, "Computational complexity in the analysis of large antennas", *Proceedings of 2004 URSI EMTS Symposium*, *Pisa*, *Italy*, pp. 141-143, 2004.

[68] J.J Bowman, T.B.A. Senior, P.L.E. Uslenghi, *Electromagnetic and acoustic scattering by simple shapes*, North Holland Publishing Company, Amsterdam, 1969

[69] J.A. Stratton, *Electromagnetic Theory*, McGraw-Hill Book Company, New York, 1941

[70] D.R. Wilton, *Computational Methods*, Department of Electrical and Computer Engineering, University of Houston

[71] O. M. Bucci, G. D'Elia, M.Santojanni, "A Fast Multipole Approach to the Scattering Evaluation based on a Non Redundant Implementation of the Method of Auxiliary Sources", scheduled for publication in *J. of E.M.Waves and Appl*., volume 20, no.13, 2006